Compound-component and conditional discrimination of colors and odors by honeybees: Further tests of a continuity model

P. A. COUVILLON and M. E. BITTERMAN
University of Hawaii, Honolulu, Hawaii

In experiments previously reported, individual honeybees were trained in a variety of problems to discriminate color-odor compounds. The results could be modeled accurately on the assumption that the associative strength of each component of a compound stimulus changes independently with reinforcement or nonreinforcement of the compound (independence rule) and that the associative strength of a compound is equal to the sum of the strengths of its components (summation rule). In the present experiments, which were designed to challenge the model, honeybees were trained in compound-component problems (choosing between compounds and their separate components) and in conditional problems (choosing between colors on the basis of a common odor or between odors on the basis of a common color). The new data, together with all of the previous data, could be modeled accurately on the further assumption that interaction among the components of a compound generates a new, compound-unique component that gains and loses associative strength in the same way as other components and contributes in the same way to the strength of the compound; the independence and summation rules continue to apply.

In recent experiments with honeybees, we studied the discriminative learning of individual foragers shuttling back and forth between the hive and the sill of an open laboratory window. On the sill, two targets were presented, one (S+) containing a drop of sucrose solution and the other (S−) containing a drop of tap water that was unacceptable to the animals. Using targets that at first differed only in odor (Couvillon & Bitterman, 1985, 1986), we were able to model performance in a variety of problems with simple equations for computing growth or decline in associative strength produced by reinforcement or nonreinforcement and for predicting choice on the basis of relative strength. Later (Couvillon & Bitterman, 1987), adding a difference in color, we were able to model performance in a variety of compound problems on the further assumption that the associative strength of each component of a compound changes independently with reinforcement or nonreinforcement of the compound (*independence rule*) and that the strength of a compound equals the summed strengths of its components (*summation rule*). In short, the parsimonious "conditioning-extinction" interpretation of Hull (1929) and Spence (1936)—or the "continuity" interpretation, as it soon came to be called (Lashley, 1942)—proved perfectly adequate to the compound data.

Change in associative strength is expressed in the model by a linear equation (Bush & Mosteller, 1951) in the familiar notation of Rescorla and Wagner (1972):

$$\Delta V_A = \beta(\lambda - V_A),$$

with V_A representing the associative strength of component A; ΔV_A, the change in strength produced by reinforcement or nonreinforcement; λ, the asymptotic associative strength; and β, the learning-rate parameter. Rescorla and Wagner's salience parameter, α, does not appear in the equation on the assumption that the colors and odors used are equally salient, supporting evidence for which is provided in our 1987 paper. Although associative strength may vary with α, the prediction of choice, which is based on relative strength, does not. The value of λ is taken as 1 when there is reinforcement (on the assumption that the magnitude of reinforcement—feeding to repletion on 50% sucrose—is maximal), as 0 when there is no reinforcement, and provision is made in the model for the possibility that the value of β is different for reinforcement and nonreinforcement ($U\beta$ = incremental, or upβ; $D\beta$ = decremental, or downβ). When choice of a compound containing A is reinforced, $\Delta V_A = U\beta(1 - V_A)$; when it is not reinforced, $\Delta V_A = D\beta(0 - V_A) = -D\beta \cdot V_A$. The independence rule is implemented by separately incrementing or decrementing each component of a reinforced or nonreinforced compound. Where A and B are two odors, X and Y two colors, and the animal is given a choice of AX and BY with AX reinforced, V_A and V_X are incremented separately if the animal chooses AX; if it goes first to BY, the associative strengths of B and Y

are decremented before those of A and X are incremented. By the summation rule, $V_{AX} = V_A + V_X$ and $V_{BY} = V_B + V_Y$.

The relative associative strength, r, of AX as compared with BY is expressed as $V_{AX}/(V_{AX}+V_{BY})$. A simple computer program evaluates r on the basis of the initial associative strengths of the components (which are assumed to be greater than 0 in consequence of pretraining with the components), dictates a choice, and then calculates the resulting changes in associative strength. Choice is dictated by one of a family of power functions that may be incorporated in the program—functions relating probability of choice, P, to r. Some examples are shown in Figure 1. To generate the choice functions, we used the scaling equations

$$P = .5 + s(2r-1)^K$$

for $r \geq .5$, and

$$P = .5 - s(1-2r)^K$$

for $r < .5$, which, with $s = .5$, yield a set of sigmoid curves passing through the points $(0,0)$, $(.5,.5)$, and $(1,1)$. When $K > 1$, P changes slowly in the region of $r=.5$ and progressively faster at the extremes; when $K < 1$, the opposite is true; and when $K = 1$, the relation between P and r is linear. (The curves of Figure 1 are plotted only for $r \geq .5$, because all the functions are perfectly symmetrical for $r < .5$.) When $s > .5$, with P truncated to $0 \leq P \leq 1$, the functions are steeper; the truncation implies, of course, that the choice measure is insensitive to differences in extreme values of r. For each of a variety of functions, we set out to find the initial value of $V_A = V_B = V_X = V_Y$ (the value after pretraining) and the values of $U\beta$ and $D\beta$ that would yield the minimal root-mean-square (RMS) deviation of predicted from obtained

choices in the entire series of compound problems—that is, the difference between the computed and the actual proportions of animals making a given choice on each visit—and we compared the functions in terms of their minimal RMS deviations. Good fits were achieved with the cluster of five similar functions plotted in the upper region of Figure 1, among which $K=.75$, $s=.833$ and $K=.5$, $s=.625$ were the best. For both functions, the RMS deviation was minimal (.114 or .115) with initial $V_A = V_B = V_X = V_Y = .4$, $U\beta = .15$, and $D\beta = .25$, reflecting the rapid changes with reinforcement and nonreinforcement that were observed in the performance of the animals.

As noted in our 1987 paper, the good fits to the compound data were surprising, because the results of some of our earlier work on compound conditioning with resistance to extinction as the measure of associative strength (Couvillon & Bitterman, 1982) seemed to call into question both the independence and the summation rules. For example, the earlier work produced clear evidence of overshadowing, which suggested that the components of a compound stimulus might in some sense compete with each other for associative strength, and clear evidence as well of compound uniqueness, which suggested that a compound stimulus might be more than a loose aggregation of components. We began the analysis of our 1987 results with the simple model in the expectation that it would fail, hoping only to be guided by the nature of the failure (the regions of the data to which it seemed especially inadequate) in the development of a better model. Concluding from the good fits obtained that the various problems chosen did not after all sufficiently constrain the model, we turned in the present work to more demanding ones, emphasizing to begin with the question of compound uniqueness.

EXPERIMENT 1:
COMPOUND–COMPONENT DISCRIMINATION

Pavlov (1927) reported that his dogs could be trained to discriminate between a compound and its components, whether the compound was reinforced and the components separately nonreinforced (positive patterning) or the reverse (negative patterning), and comparable results have since been obtained by other techniques in work with other animals (Bellingham, Gillette-Bellingham, & Kehoe, 1985; Rescorla, 1972; Whitlow & Wagner, 1972). If, in our choice situation, the positive-patterning problem is implemented as $AX+$ versus $A-$ on half the trials and $AX+$ versus $X-$ on the rest (A is an odor and X is a color) and the negative-patterning problem as $A+$ versus $AX-$ on half the trials and $X+$ versus $AX-$ on the rest, the model predicts better-than-chance performance in the positive-patterning problem and poorer-than-chance performance in the negative-patterning problem (by the summation rule, V_{AX} must be greater than V_A or V_X). Before considering how the model would have to be changed to
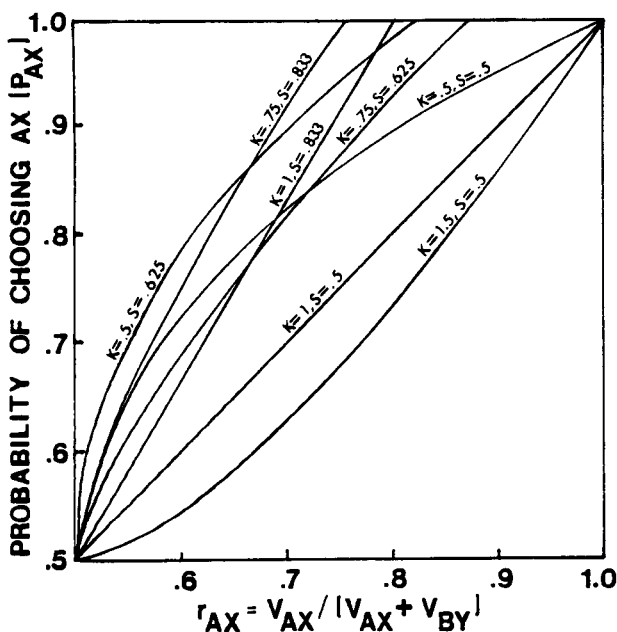

Figure 1. Choice functions.

account for negative patterning, it may be well to look at the actual performance of honeybees both in positive and in negative problems.

Method

Subjects. The subjects were 32 honeybees (*Apis mellifera*) foraging for nectar. They came from our own hives, situated in the vicinity of the laboratory, and all were experimentally naive.

Procedure. Individual bees were pretrained to fly from their hives to the laboratory and to drink to repletion from a large drop (about 100 μl) of 50% sucrose solution on a target that was set on the sill of an open window. An animal was selected at random from a group of foragers at a feeding platform equipped with a large jar of 10%-12% sucrose solution, carried in a small matchbox to the laboratory, and placed on the target. There it was permitted to drink its fill of 50% solution (during which time it was marked with a spot of colored lacquer) and then to fly to the hive. Typically, the animal returned to the laboratory of its own accord in 3-4 min, continuing thereafter to shuttle back and forth between the hive and the laboratory as long as food was available. If the marked bee did not return to the laboratory after the first placement, it was picked up again at the feeding platform (where it usually could be found), carried back to the laboratory, and set down on the target once more. The pretraining ended with the subject's second return to the laboratory of its own accord. It should be noted that the work could be done with individual foragers because recruitment is not a problem under our conditions. Intruders (bees which have not themselves been carried to the laboratory) are rare and can quickly be captured, as is each subject at the conclusion of its training.

The targets were covered petri dishes of clear plastic, 5.5 cm in diameter, with gray covers. In each cover, eight equally spaced holes, .5 cm in diameter, were drilled at the outer circumference. To some of the covers, rings of green plastic, 1 cm wide and 5.5 cm in outside diameter, were cemented, with holes in the rings corresponding to the holes in the covers. The dishes themselves contained wads of cotton that were either impregnated with the odor of peppermint or unimpregnated. In all, three sets of targets were used in the training: green ring with peppermint scent (*AX*), peppermint scent without the green ring (*A*), and green ring without the peppermint scent (*X*). The covers used on each visit were washed and exchanged for others in the same set after the visit in order to balance extraneous stimuli. For purposes of pretraining only, there was an additional set of targets without either the green ring or the peppermint scent.

On each of 32 training visits, the animal found two targets placed 10 cm apart on a line parallel to the outer edge of the window sill. One of the targets contained a drop of 50% sucrose solution, and the other contained a drop of tap water that was distinguishable from the sucrose only by taste. The target first chosen on each visit was recorded, and the trial ended when the animal had found the correct target, fed to repletion, and gone back to the hive. For 16 animals in the positive patterning group (P), there were 16 visits to *AX*+*A*− intermixed in quasi-random sequence with 16 visits to *AX*+*X*−. The sequence began with *A*− for half the animals and with *X*− for the rest. Position also was balanced: *AX*+ was at left on half the *A*− visits and on half the *X*− visits, and at right on the remaining visits, again in quasi-random sequence. In all, four different training orders were used, each for 4 animals in the group. For 16 animals in the negative patterning group (N), the training was exactly the same as for Group P except that the components rather than the compound were reinforced (*A*+*AX*− and *X*+*AX*−).

Results

The performance of the two groups is plotted in Figure 2 in terms of the proportion of animals in each group choosing correctly on each visit. The two curves are very much the same except on the first visit, when both groups showed a preference for the component over the compound, which tended to impair performance in Group P and elevate it in Group N. Thereafter, the curves merge, gradually approaching a common asymptote of about 80% correct choice, and the overall performance of the two groups does not differ significantly (median test, Fisher's exact $p < .14$). Although the asymptotic level is not high, each of the 32 animals made fewer than 16 errors in the 32 training visits (exact $p < .0001$).

One way to account for the better-than-chance performance of Group N is to assume "afferent neural interaction" (Hull, 1943). In Hull's view, the animals discriminate not between *AX* and the components but between *A'X'* and the components, where *A'* and *X'* are the components as altered by interaction. (Actually, of course, it is not the component stimuli that are assumed to be altered by interaction but the corresponding afferent components, which are distinguished in Hull's notation.) This solution requires the addition of a generalization principle to the model, which ultimately will be necessary in any case. Another proposal (Rescorla, 1972; Whitlow & Wagner, 1972) is that, instead of changing *A* and *X*, the interaction generates a compound-unique component, which we may refer to here as *Q*. In this view, the animals distinguish *AXQ* from *A* and *X* (again conceived as affer-
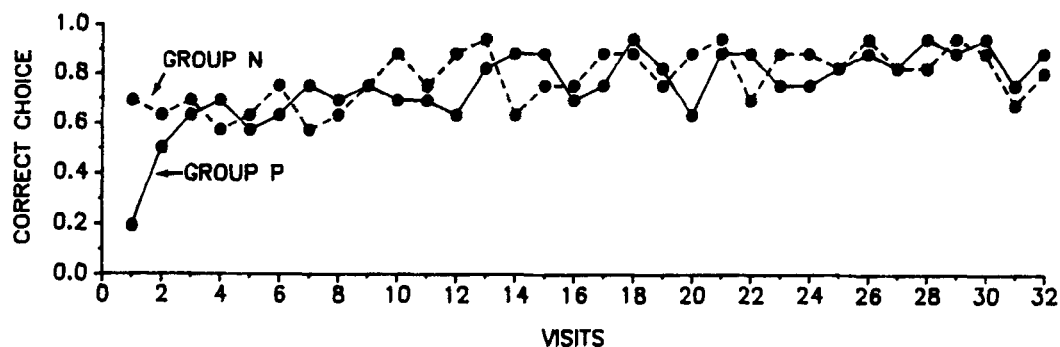

Figure 2. Performance in the positive-patterning (Group P) and negative-patterning (Group N) problems of Experiment 1.

ent components), which helps us to account for negative patterning only if the model is extended to permit Q to become inhibitory. We have been able thus far to deal with all of our data on the assumption that the function of nonreinforcement is simply to reduce associative strength (excitation), which cannot be less than zero. It is interesting to note that neither proposal requires us to abandon the independence and summation rules, although they now apply to afferent components rather than to stimulus components.

While considering these alternatives, we hit upon another way of looking at the patterning problems that requires neither of them. The idea is to think of A not as an odor in the absence of a color (peppermint in the absence of the green ring) but as an odor in the presence of a different color, Y (the plain gray target), and to think of X not as a color in the absence of an odor (green ring in the absence of peppermint) but as a color in the presence of a different odor, B (the odor of the target with unimpregnated cotton). If this analysis is correct, the animals discriminate AX from AY and BX, and we should be able to simulate their performance with no change at all in the model. In the positive case, with only A and X acquiring any associative strength, V_{AX} must be greater than V_{AY} or V_{BX}; in the negative case, with B and Y consistently reinforced but A and X only inconsistently reinforced, V_{AX} must be less than V_{AY} or V_{BX}. Even before a formal simulation, however, it should be noted that the analysis provides an explanation of the clear preference for the "component" over the "compound" shown by the animals on the first training visit. Since the pretraining was to the plain gray target filled with unimpregnated cotton (BY), the animals should prefer either AY or BX to AX.

The simulation was done with the same parameters—the same choice function ($K=.75$, $s=.833$), the same initial V values of pretrained components $V_B = V_Y = .4$), and the same learning rates ($U\beta = .15$, $D\beta = .25$)—that yielded the best fit to the combined data of our previous compound experiments (Couvillon & Bitterman, 1987). Since we were, for the moment, interested only in getting some general sense of the validity of the analysis, we made the simplifying assumption that all the components were equal in salience (the salience of the unimpregnated cotton equal to that of the scented cotton and the salience of the plain gray target equal to that of the gray target with green ring). Nevertheless, the results for two groups of 100 stat-bees were close enough to those for the real bees, showing even the clear preference for "components" on the first visit, that we were encouraged to repeat the work with the standard stimuli of our 1987 experiments, which are known to be equally salient and discriminable and are described below.

EXPERIMENT 2:
FURTHER STUDY OF
"COMPOUND-COMPONENT" DISCRIMINATION

If compound-component problems as implemented in Experiment 1 in effect only require the animals to dis-

criminate AX from AY and BX, we should get very much the same results in explicit training on such problems with the stimuli used in our previous work on compound discrimination—homogeneous green and blue targets instead of green circle on gray and gray alone as X and Y, and the scents of peppermint and geraniol instead of peppermint and unimpregnated cotton as A and B. With these stimuli, moreover, quantitative prediction of the new results was possible. For the prediction, we again used the choice function ($K=.75$, $s=.833$), the initial V values for A, B, X, and Y ($=.4$), and the learning rates ($U\beta=.15$, $D\beta=.25$) that had yielded the best fit to the data of our previous compound experiments. A new simulation nevertheless was necessary here because of a change in the pretraining procedure; for conformity with our previous work on compound discrimination, the animals in Experiment 2 experienced all four components in pretraining.

Method
Subjects. The subjects were 32 foragers from our own hives, all experimentally naive.

Procedure. The pretraining procedure was the same as in Experiment 1, except that it was done with a set of targets that were half-blue, half-green, and scented both with peppermint and geraniol. In the discriminative training, four sets of targets were used—green-peppermint, green-geraniol, blue-peppermint, and blue-geraniol—but only three of them for any given subject. For example, 4 of 16 positive-patterning animals (Group P) were trained with green-peppermint (AX) reinforced versus blue-peppermint (AY) or green-geraniol (BX) nonreinforced; 4 of 16 negative-patterning animals (Group N) were trained with the same targets but with blue-peppermint and green-geraniol reinforced and green-peppermint nonreinforced. The two pairs of targets (AX,AY and AX,BX) appeared equally often in two balanced, quasi-random sequences, with each target of each pair presented equally often at right and left. Other P and N subgroups were trained with AY versus AX and BY; BX versus BY and AX; and BY versus BX and AY. Since the performance of the animals of Experiment 1 appeared to be asymptotic after 16 training visits, that was the number of visits permitted the animals of the present experiment.

Results
In Figure 3, the performance of the two groups is plotted in terms of the proportion of animals (Group P, upper panel, and Group N, lower panel; large circles) choosing correctly on each visit. The marked initial separation of the curves of Figure 2 does not appear in these curves and, because of the pretraining with all four components, is not, of course, predicted by the theory. There is some indication of better overall performance in Group P than in Group N, but the difference is not statistically reliable (median test, exact $p > .07$). What appears to be the asymptotic level of performance again is not high, but there is no question about better-than-chance performance; 31 of the 32 animals made fewer than eight errors in the 16 visits (exact $p < .0001$).

The predicted results, based on simulations with 100 stat-bees, also are plotted in Figure 3 (small circles). Although the fit is encouraging, the performance of the real animals is somewhat underpredicted by the model. If we had no other information, we might be tempted to
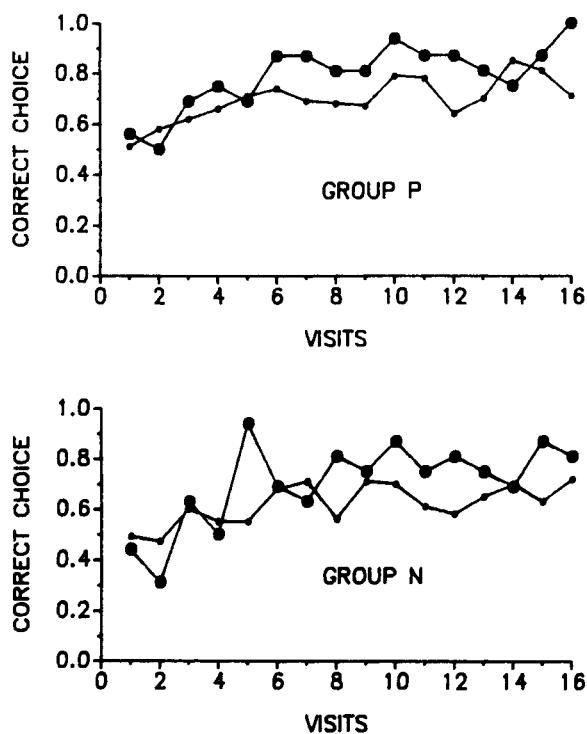
Figure 3. Actual (large circles) and simulated (small circles) performance of Groups P and N in Experiment 2. The simulation was done with the parameters that yielded the best fit to the Couvillon and Bitterman (1987) data.

treat the discrepancy simply as a parametric problem, but the results of a further experiment, to which we now turn, showed that a change in the theory was required.

EXPERIMENT 3: CONDITIONAL DISCRIMINATION

In early experiments, Lashley (1938) trained rats to choose between upright and inverted triangles on black or striated backgrounds, with the upright triangle correct when the backgrounds were black and the inverted triangle correct when the backgrounds were striated. Lashley's results were confirmed by North, Maller, and Hughes (1958), whose further tests showed that the performance of the animals was determined by properties of the separate compounds as distinct from properties of the pairs of compounds ("configurational" properties). Conditional discrimination has been demonstrated also in rabbits (Saavedra, 1975) and goldfish (Bitterman, 1984) trained with compounds of visual and auditory stimuli. In the present experiment with honeybees, we used two different conditional problems, one that required the animals to choose between differently colored targets on the basis of a common odor, and one that required them to choose between differently scented targets on the basis of a common color. The model predicted failure, of course: Trained, for example, with $AX+AY-$ and $BY+BX-$, an animal could prefer the positive compound

of each pair only if $V_A+V_X > V_A+V_Y$ and $V_B+V_Y > V_B+V_X$, or, by addition, $V_A+V_B+V_X+V_Y > V_A+V_B+V_X+V_Y$. Nevertheless, we had reason to look at least for better-than-chance performance on the basis of an earlier demonstration of color–odor compound uniqueness as a determinant of resistance to extinction in honeybees (Couvillon & Bitterman, 1982). In that experiment (which was done with other stimuli), bees that had been fed by turns on AX and BY responded more persistently to the two compounds than to AY and BX.

Method

Subjects. The subjects were 24 foragers from our own hives. All were experimentally naive.

Procedure. The pretraining procedure was the same as in Experiment 2, with special targets half-green, half-blue, and scented both with geraniol and with peppermint. There were 32 training visits. The 12 subjects of the color group were trained to choose between green and blue targets that were scented with geraniol on half the visits and with peppermint on the rest, in quasi-random order. The 12 subjects of the odor group were trained to choose between geraniol- and peppermint-scented targets that were green on half the visits and blue on the rest, in quasi-random order. Green-geraniol and blue-peppermint were correct for half the animals in each group, and blue-geraniol and green-peppermint were correct for the rest. Each compound was presented half the time on the left and half the time on the right, again in quasi-random order.

Results

In Figure 4, the performance of the two groups is plotted in terms of the proportion of animals in each group choosing correctly on each visit. The curves suggest that the problems are equivalent in difficulty (median test, exact $p < .34$) and certainly not insoluble, although the asymptotic performance is far from perfect. Of the 24 animals, only 2 failed to show better-than-chance performance in the 32 visits, 1 of them making 17 errors and the other 16 (exact $p < .0016$). In the last half of the training (the last 16 visits), the mean number of incorrect choices was 3.8 (76% correct; $SE=.48$). Given this performance, we seemed to have no alternative but to incorporate in the model some notion of compound uniqueness, and of the two suggested in the vertebrate literature we chose to rely on the more easily implemented—the assumption of compound-unique components. The implementation is easier because, for the moment at least, we can ignore not only generalization (which the Hullian interpretation would require us to deal with immediately) but also (given our analysis of the negative patterning in Experiment 1) the possibility of inhibition.

Consider again the conditional problem involving $AX+AY-$ and $BY+BX-$: If A and X together generate a new element, Q, and if B and Y together generate a new element, R, then, by the independence rule, a rewarded response to AX increments V_Q as well as V_A and V_X, and a rewarded response to BY increments V_R as well as V_B and V_Y. By the summation rule, $V_{AX} = V_A+V_X+V_Q$, and $V_{BY} = V_B+V_Y+V_R$. Note that V_Q and V_R are never decremented because AX and BY always are reinforced.
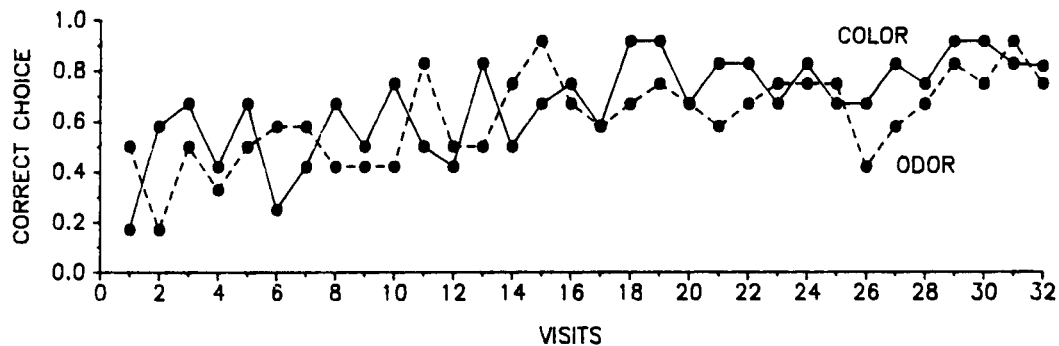
Figure 4. Performance of the color and odor groups in the conditional problems of Experiment 3.

Note, too, that the hypothetical elements unique to AY and BX (S and T) can be ignored because AY and BX are never reinforced; $V_S = V_T = 0$. Given the implication of our previous results (Couvillon & Bitterman, 1987) that the colors and odors used are equally salient ($\alpha_A = \alpha_B = \alpha_X = \alpha_Y = 1$), the model suggests that the four possible conditional problems ($AX+AY-,BY+BY-$; $AY+AX-,BX+BY-$; and $AX+BX-,BY+AY-$; and $BX+AX-,AY+BY-$) are equivalent if Q, R, S, and T are assumed to be equally salient, which seems a reasonable first approximation. There is, however, no way to decide in advance what the salience of the compound-unique components should be, except that it must be greater than zero.

Our procedure, therefore, was to simulate the performance of 100 bees in the conditional problem, with the salience of the compound-unique components varying from 0 to 1 in steps of .1, and to compare the results with the combined performance of the color and odor animals. Again, we used the same choice function ($K = .75$, $s = .833$), initial V values of the pretrained components ($= .4$), and learning rates ($U\beta = .15$, $D\beta = .25$) that we used throughout (those that yielded the best fits to the combined data of our previous compound experiments). As expected, the fit was poor for $\alpha = 0$ (RMS deviation of simulated from obtained values $= .25$). With increases in the assumed salience of the compound-unique components, the RMS deviation decreased progressively, reaching a minimum of .098 at $\alpha = .7$, and then increased again with further increases in α to .117 for $\alpha = 1$.

A question that now must be considered is why, if compound-unique components of substantial salience are generated in compound problems, we should have been able to simulate our 1987 results so well without taking them into account. One possibility is that compound-unique components acquire control of performance only when other components are nondifferentially reinforced (Spence, 1952), which would violate the independence rule. A second possibility is that compound-unique components always function as do other components, but that the error of neglecting them in our previous simulations was compensated for by complementary errors in the

selection of a choice function and the estimation of learning rates.

To examine the second (much simpler) possibility, we simulated performance in all of the compound problems described in our 1987 paper together with performance in the present Experiments 2 and 3, using each of the cluster of five choice functions shown in the upper region of Figure 1 that previously had provided the best fits of the first compound model to the 1987 data. We also used $K = 1$, $s = .5$, the linear function that had provided a much poorer fit, and—entering again a region of the functional space unexplored since the beginning of our work with compounds—the still more different $K = 1.5$, $s = .5$. These functions, too, are shown in Figure 1. For each function, performance in all problems was simulated with initial values for $V_A = V_B = V_X = V_Y$ of .1, .2, .3, or .4, values for $U\beta$ and $D\beta$ (varied factorially) of .15, .20, .25, .30, or .35, and saliences for the compound-unique components (assumed to be equal) ranging from .1 to 1 in steps of .1; the initial V values of the compound-unique components were taken as 0, since they were not encountered in pretraining. In all, there were 1,000 simulations for each choice function (4 initial V values $\times$ 5 values of $U\beta$ $\times$ 5 values of $D\beta$ $\times$ 10 values of α), each simulation for 100 stat-bees. The RMS computation for each simulation was based on the 108 data points provided by the earlier compound experiments and 61 new points provided by the present Experiments 2 and 3. The 16-visit positive- and negative-patterning problems of Experiment 2 were treated as separate problems, because there was no a priori reason to suppose that they should be equivalent. Together, they contributed 30 of the 61 new points; the first visit in each case was, as usual, excluded on the ground that any deviation from .5, either in the simulated or in the actual performance, must represent random error. The 32-visit conditional discrimination of Experiment 3 contributed 31 of the new points; the color and odor problems were not treated separately because they are mathematically equivalent on the assumption that the colors and odors used are equally salient (Couvillon & Bitterman, 1987). Actually, there were two RMS computations for each simulation, one based on untransformed
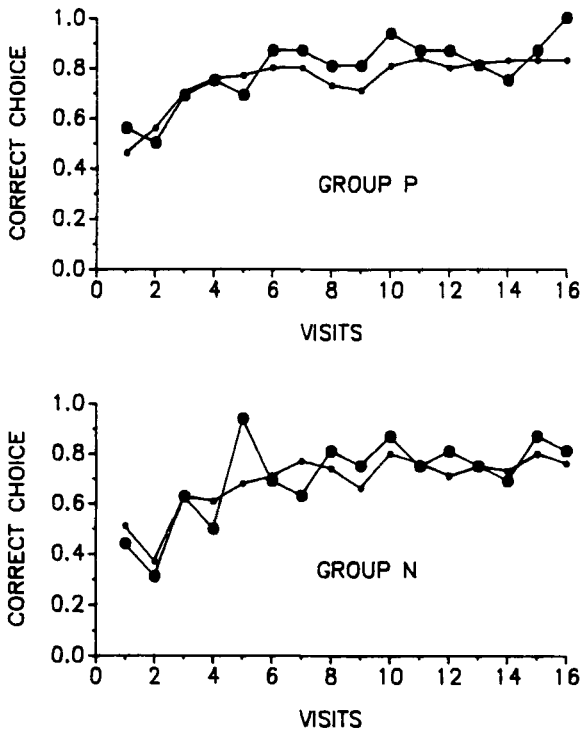
Figure 5. Actual (large circles) and simulated (small circles) performance of Groups P and N in Experiment 2. The simulation was done with the parameters that yielded the best fit to the combined data of the present and previous (Couvillon & Bitterman, 1987) experiments.

proportions and another after arcsin transformation, each proportion, P, being replaced by an angle, ϕ, measured in degrees, according to the equation

$$\phi = (180/\pi)\sin^{-1}P^{.5}.$$

The RMS scores based on untransformed proportions have more intuitive meaning, but (because the variance of a proportion decreases with distance from .5) they may not give sufficient weight to small deviations at the extremes of the scale. As it happens, the pattern of results proved to be the same for both sets of scores.

The cluster of five choice functions (Figure 1) that had yielded good fits of the first compound model to the 1987

data continued to yield good fits of the revised model to the entire set of compound data. Here, however, the fits were best not as before for $K=.5$, $s=.625$ and $K=.75$, $s=.833$, the two steepest functions (untransformed RMS = .125, arcsin RMS = 7.2°), but for $K=.75$, $s=.625$ and $K=1$, $s=.833$, two of the less steep functions (untransformed RMS = .119, arcsin RMS = 6.8°). The goodness of fit for $K=.75$, $s=.625$ (initial $V=.3$, $U\beta=.2$, $D\beta=.35$, and $\alpha=.7$) is illustrated in Figure 5 with the data of the two groups of Experiment 2 and in Figure 6 with the pooled conditional data of Experiment 3 (large circles, obtained proportions of correct choice; small circles, simulated proportions). For $K=.5$, $s=.5$, the goodness of fit declined again (untransformed RMS = .123, arcsin RMS = 7.1°). As expected from previous work, the best fit possible with the "matching" function ($K=1$, $s=.5$) was considerably poorer (untransformed RMS = .139, arcsin RMS = 8.3°) than for any of the clustered five, and the best fit possible with the more extreme $K=1.5$, $s=.5$ was poorer still (untransformed RMS = .171, arcsin RMS = 10.2°).

It may be well to note again that the RMS deviation of simulated from obtained values is a reflection not only of model error (σ_m) but also of random error in the data (σ_d) and of random error in the simulation process itself (σ_s). To estimate the simulation error, we did a second simulation of performance in the entire set of compound problems for 100 stat-bees with $K=.75$, $s=.625$, initial $V=.3$, $U\beta=.2$, $D\beta=.35$, and $\alpha=.7$, the parameters that yielded the fit illustrated in Figures 5 and 6. On the basis of the RMS deviation of the two sets of simulated values from each other, which turned out to be .056, σ_s may be estimated from the equation

$$\sigma_s^2 + \sigma_s^2 = (.056)^2$$

as .040. On the assumption of independence, it is possible also to estimate the data and model error combined (σ_{dm}) from the equation

$$\sigma_s^2 + \sigma_{dm}^2 = (.119)^2$$

as .112. It is reasonable to assume from the nature of the data that the model error alone is rather small.
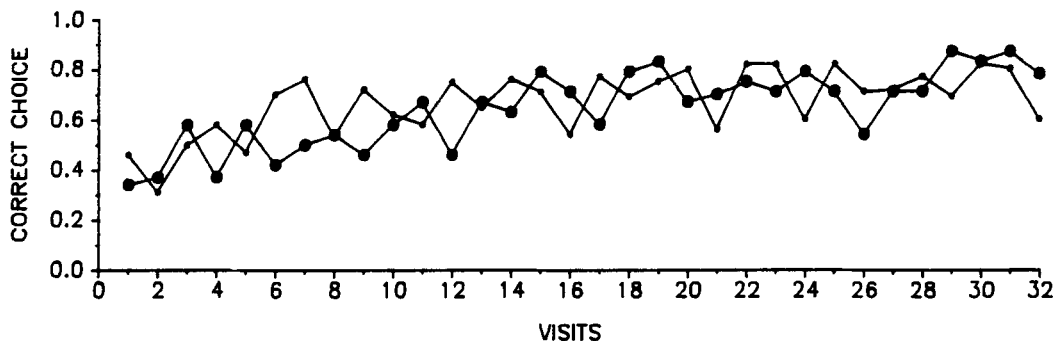

Figure 6. Actual (large circles) and simulated (small circles) performance of the combined color and odor groups in the conditional problems of Experiment 3. The simulation was done with the parameters that yielded the best fit to the combined data of the present and previous (Couvillon & Bitterman, 1987) experiments.

DISCUSSION

We came to this work with a simple continuity model that nicely accounted for the choices made by honeybees in a wide range of simultaneous color-odor discrimination problems but that was not, we suspected, sufficiently constrained by those problems. Our skepticism was based on the results of prior conditioning experiments with resistance to extinction as the measure of associative strength, in which a variety of phenomena were discovered that seemed to challenge the independence and summation rules. Of those phenomena, we turned first to compound uniqueness, further evidence of which was sought in a new series of choice problems culminating in a demonstration of conditional discrimination. The new data could not be accounted for in terms of the existing model, but, together with all of the previous data, they could be simulated accurately on the further (purely perceptual) assumption that interaction among the components of a compound stimulus generates a distinctive (compound-unique) component that changes in associative strength and contributes to the strength of the compound in the same way as other components. The extended model remains a continuity model in its adherence to the parsimonious independence and summation rules.

Still waiting to be considered are other phenomena discovered in the extinction experiments that seem to present more fundamental challenges to the model, among them overshadowing and potentiation (Couvillon & Bitterman, 1982). Evidence against the independence rule may not have appeared in the 1987 choice experiments for any of a number of reasons, again perhaps because the designs were not sufficiently constraining. The possibility also must be considered that the training technique itself, however efficient, is relatively insensitive; simultaneous discrimination experiments with vertebrates certainly did not in their heyday, now long gone, produce very much in the way of clear and replicable evidence of noncontinuity (Sutherland & Mackintosh, 1971). The model is challenged as well by indications of inhibitory control that appeared in the extinction experiments (Couvillon & Bitterman, 1980) but may prove difficult to find in choice training if simultaneous discrimination experiments with vertebrates are any guide; although inhibition often has been inferred, as, for example, to account for transposition (Spence, 1937), other interpretations are available and more direct evidence is wanting (Macphail, 1972). We plan nevertheless to continue to test the limits of the honeybee choice model along the lines suggested by the extinction experiments.

REFERENCES

BELLINGHAM, W. P., GILLETTE-BELLINGHAM, K., & KEHOE, E. J. (1985). Summation and configuration in patterning schedules with rat and rabbit. *Animal Learning & Behavior*, 13, 152-164.

BITTERMAN, M. E. (1984). Migration and learning in fishes. In J. D. McCleave, G. P. Arnold, J. J. Dodson, & W. H. Neill (Eds.), *Mechanisms of migration in fishes* (pp. 397-420). New York: Plenum.

BUSH, R. R., & MOSTELLER, F. (1951). A methematical model for simple learning. *Psychological Review*, 58, 313-323.

COUVILLON, P. A., & BITTERMAN, M. E. (1980). Some phenomena of associative learning in honeybees. *Journal of Comparative & Physiological Psychology*, 94, 878-885.

COUVILLON, P. A., & BITTERMAN, M. E. (1982). Compound conditioning in honeybees. *Journal of Comparative & Physiological Psychology*, 96, 192-199.

COUVILLON, P. A., & BITTERMAN, M. E. (1985). Analysis of choice in honeybees. *Animal Learning & Behavior*, 13, 246-252.

COUVILLON, P. A., & BITTERMAN, M. E. (1986). Performance of honeybees in reversal and ambiguous-cue problems: Tests of a choice model. *Animal Learning & Behavior*, 14, 225-231.

COUVILLON, P. A., & BITTERMAN, M. E. (1987). Discrimination of color-odor compounds by honeybees: Tests of a continuity model. *Animal Learning & Behavior*, 15, 218-227.

HULL, C. L. (1929). A functional interpretation of the conditioned reflex. *Psychological Review*, 36, 498-511.

HULL, C. L. (1943). *Principles of behavior*. New York: Appleton-Century-Crofts.

LASHLEY, K. S. (1938). Conditional reactions in the rat. *Journal of Psychology*, 6, 311-324.

LASHLEY, K. S. (1942). An examination of the continuity theory as applied to discriminative learning. *Journal of General Psychology*, 26, 241-265.

MACPHAIL, E. M. (1972). Inhibition in the acquisition and reversal of simultaneous discriminations. In R. A. Boakes & M. S. Halliday (Eds.), *Inhibition and learning* (pp. 121-151). London: Academic Press.

NORTH, A. J., MALLER, O., & HUGHES, C. (1958). Conditional discrimination and stimulus patterning. *Journal of Comparative & Physiological Psychology*, 51, 711-715.

PAVLOV, I. P. (1927). *Conditioned reflexes*. Oxford: Oxford University Press.

RESCORLA, R. A. (1972). "Configural" conditioning in discrete-trial bar pressing. *Journal of Comparative & Physiological Psychology*, 79, 307-317.

RESCORLA, R. A., & WAGNER, A. R. (1972). A theory of classical 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.

SAAVEDRA, M. A. (1975). Pavlovian compound conditioning in the rabbit. *Learning & Motivation*, 6, 314-326.

SPENCE, K. W. (1936). The nature of discrimination learning in animals. *Psychological Review*, 43, 427-449.

SPENCE, K. W. (1937). The differential response in animals to stimuli varying within a single dimension. *Psychological Review*, 44, 430-444.

SPENCE, K. W. (1952). The nature of response in discrimination learning. *Psychological Review*, 59, 89-93.

SUTHERLAND, N. S., & MACKINTOSH, N. J. (1971). *Mechanisms of animal discrimination learning*. New York: Academic Press.

WHITLOW, J. W., JR., & WAGNER, A. R. (1972). Negative patterning in classical conditioning: Summation of response tendencies to isolable and configural components. *Psychonomic Science*, 27, 299-301.