View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Universidade do Minho: RepositoriUM

Title: RESPONDING BY EXCLUSION IN TEMPORAL DISCRIMINATION TASKS

Authors: Nathalia Sabaine Cippola[1], Camila Domeniconi[1], Armando Machado[2]

1. **Federal University of São Carlos, Brazil**
2. **University of Minho, Portugal**

Correspondence Address

Emails: armandom@psi.uminho.pt
camila@ufscar.br

ABSTRACT

Responding by exclusion, one of the most robust phenomena in Experimental Psychology, consists of choosing an undefined comparison stimulus given an undefined sample, when the comparison stimulus is presented next to other experimentally defined stimuli. The goal of the present study was to determine whether responding by exclusion could be obtained using samples that varied along a single dimension. Using a double temporal bisection task, ten university students learned to choose visual comparisons (colored circles) based on the duration of a tone. In tests of exclusion, sample stimuli with new durations were followed by comparison sets that included one previously trained, defined comparison (colored circle) and one previously untrained, undefined comparison (geometric shape). Subjects preferred the defined comparisons following the defined samples and the undefined comparisons following the undefined samples, the choice pattern typical of responding by exclusion. The use of samples varying along a single dimension allows us to study the interaction between stimulus generalization gradients and exclusion in the control of conditional responding.

Keywords: responding by exclusion; temporal stimuli; stimulus control; conditional relation; university students.

INTRODUCTION

The term "responding by exclusion" describes a particular form of responding observed in matching-to-sample tasks. Suppose that, when exposed to comparison stimuli C_1 and C_2 after the presentation of the sample stimulus M_1 or M_2, a subject learns to choose C_1 after M_1 and C_2 after M_2. We represent these conditional discriminations as $M_1 \rightarrow C_1$ and $M_2 \rightarrow C_2$. Subsequently, after the presentation of a sample stimulus M_X, a sample not previously related to any comparison stimulus, the subject is given a choice between C_1 and C_Y or $C_2$and C_Y, in which the comparison C_Y is unrelated to any previous sample. The sample M_X and the comparison C_Y are both undefined stimuli. These relationships can be represented as follows

$$\text{Training: } \begin{cases} M_1 \rightarrow C_1 \\ M_2 \rightarrow C_2 \end{cases} \text{; Testing: } \begin{cases} M_X : C_1 \text{ or } C_Y \\ M_X : C_2 \text{ or } C_Y \end{cases}.$$

Several studies have shown that, given the undefined sample M_X, the subject tends to exclude C_1 or C_2, the comparisons previously related to samples other than the current sample, and choose C_Y, the comparison that also is undefined. The preference for C_Y given M_x occurs without any explicit training or teaching of the $M_X \rightarrow C_Y$ relation. In addition to responding by exclusion, this response pattern has been termed "rapid mapping" by psycholinguists and "symbolic mapping" by behavioral analysts (e.g., Bates, Benigni, Bretherton, Camaioni, &Volterra, 1979; Carey & Bartlett, 1978; Costa, Wilkinson, McIlvane, & de Souza,2001; Dixon, 1977; Floor & Akhtar, 2006; Halberda, 2003; McIlvane, Kledaras, Lowry & Stoddard,1992; Pruden, Hirsh-Pasek, Golinkoff, & Hennon, 2006; Stromer, & Osborne, 1982; Wilkinson Rosenquist & McIlvane, 2009).

Responding by exclusion was first observed in a study on the teaching of conditional discriminations to adolescents with mental retardation (Dixon, 1977). To study the conditional control that a verbal stimulus exerts on the subject's choice, Dixon programmed training trials that included Greek letters or Japanese characters as comparisons and the corresponding dictated words as samples. To illustrate, on some trials, Dixon dictated the word "pi" and then presented two visual comparisons, the Greek letter π, the correct choice or S^+, and another letter that could be either θ or Y, the incorrect choice or S^-. Once the relation "pi"$\rightarrow\pi$ was learned, exclusion tests were performed. A new, undefined dictated word such as "theta" was presented as sample followed by two visual comparisons, π and θ.The subjects systematically chose the undefined visual stimulus θafter the new sample "theta", but continued to choose the defined visual stimulus πafter the defined sample "Pi". Similarly, following the undefined sample "epsilon", the subjects chose the undefined comparison Y over the defined comparison π, but continued to choose π following the defined sample "Pi". Based on these findings, Dixon suggested that the subjects responded to the new comparisons "by exclusion": When the dictated word was unknown or undefined, the subjects excluded the defined or trained visual stimulus and chose the undefined, untrained visual stimulus.

Since then several studies have attested to the robustness of responding by exclusion in subjects with typical and atypical development (Dixon, Dixon, & Spradlin, 1983; McIlvane et al., 1992; McIlvane, Munson, & Stoddard, 1988; McIlvane & Stoddard, 1981; Stromer, 1986, 1989), in children of different ages (Costa et al., 2001; Domeniconi, Costa, de Souza, & de Rose, 2006; Ferrari, de Rose, & McIlvane, 1993),and in university students (McIlvane, Kledaras, Munson, King, de Rose, & Stoddard, 1987). Other studies have suggested that some animals also respond by exclusion, including sea lions and dolphins (Kastak & Schusterman 2002; Herman et al., 1984; Schusterman & Krieger, 1984), dogs (Aust, Range, Steurer, & Huber, 2008; Kaminski, Call, & Fischer, 2004), and chimpanzees (Beran & Washburn, 2002).

However, results obtained with pigeons have been inconclusive. Cumming and Berryman (1961) trained pigeons in an identity conditional discrimination task with three pairs of colors, red, blue, and green, and then, after the discrimination was learned, replaced the blue pair by a yellow pair. On test trials, the new sample (yellow) was followed by two comparisons, one previously trained (red or green) and one untrained (yellow). The pigeons responded to the yellow stimulus at chance level. The authors concluded that pigeons, although not preferring the new stimulus, did not show neophobia because, in this case, the percentage of responses to the new stimulus would be lower than chance.

According to Clement and Zentall (2003, 2000), the results reported by Cumming and Berryman (1961) could have been due to the introduction of new stimuli during testing, a procedural feature that may have made performance by exclusion difficult to obtain. To test the idea, he authors designed a procedure in which responding by exclusion could be evaluated with no new stimuli added at testing. In one experiment, Clement and Zentall (2003) trained arbitrary conditional relations with colors as samples and shapes as comparisons. During Phase 1(pretraining), the pigeons were trained on a conditional discrimination with Red and Green both as samples and comparisons (Red-Green; Green-Red). During Phase 2, the pigeons received a conditional discriminations training that involved colors as samples and shapes as comparisons (Red-Circle; White-Plus). Finally, during the test phase the pigeons received the Green and Red samples and the Circle and Plus comparisons. The target trial involved the Green sample, which was not presented during training. If pigeons respond by exclusion, the authors reasoned, they should exclude Circle because it was paired with Red and therefore choose Plus. Results were consistent with the authors' prediction. However, it is still unclear whether the pigeons chose the Plus comparison given the Green sample reliably since the first test trials, or learned to do so only during testing (Schloegl, Bugnyar, &Aust, 2009). Moreover, it remains to be seen whether the procedural features introduced by Clement and Zentall to eliminate the potential disruptive effects of stimulus novelty, namely, the use of a familiar sample and a familiar comparison during the critical test, may not be incompatible with the inference that the pigeons responded by exclusion.

Despite the uncertainty about pigeons' performance and whether it qualifies as responding by exclusion, the results of previous studies, particularly with humans, suggest that the pattern of responding by exclusion is one of the most reliable phenomena in experimental psychology (Wilkinson, de Souza, &McIlvane, 2000). However, the generality of responding by exclusion remains unknown. The phenomenon has been studied mostly with auditory samples and visual comparisons and only rarely with visual stimuli exclusively (e.g., Clement & Zentall, 2000, 2003; Oshiro, de Souza, & Costa, 2006). In addition, the procedure itself has not varied much across studies, and therefore little is known about the factors that are critical for observing the phenomenon.

Exceptions are the studies by Costa (2008) and Costa, de Rose, and de Souza (2010), who investigated how responding by exclusion changes when lexical cues are present during testing. Because the subjects (i.e., children and individuals with developmental delays) demonstrated responding by exclusion -- choosing an undefined stimulus C_Y over alternative, defined comparisons when an undefined word M_x was dictated --, the authors added contextual cues to the dictated word, cues for plural (e.g., "mopadi<u>s</u>" instead of "mopadi"), size (e.g.,"mopad<u>inho</u>", with the Portuguese diminutive "inho" meaning "small"), or cues for action (e.g., "mopadi<u>ando</u>", similar to the English "ing" form).The results demonstrated that adding the contextual cues during the test trials altered the pattern of responses typically reported in the literature, with a trend toward fewer choices of the undefined comparison stimulus. In this case, responding was controlled by specific properties of the complex stimulus (Costa, 2008; Costa, de Rose, and de Souza, 2010).

The author also discussed the difficulty of identifying the limits of responding by exclusion. If one thinks that any stimulus can be characterized by the specific values it has along its various dimensions, regardless of how complex these dimensions might be, then the notion of an undefined stimulus can be seen in a new light, as a matter of degree rather than all-or-none. Sample or comparison, a stimulus will be more undefined the more its values along the relevant dimensions differ from the corresponding values of the defined stimuli. Or, to put it in a different but equivalent way, the multidimensional generalization gradients of the defined stimuli set the criteria to define how "undefined" a new stimulus is. The goal of the present study was to explore this conception.

To that end, we simplified the stimulus samples so that they varied along a single dimension, duration. The defined samples corresponded to tones with distinct durations (e.g., 200 ms and 600 ms) and they were trained in the conditional relations $M_{200} \rightarrow C_1$ and $M_{600} \rightarrow C_2$. The undefined samples corresponded to tones with durations that (a) were not previously trained in any conditional relation and (b) were located at different distances from the trained values of 200 and 600 ms. We expected that this new procedure would allow us to observe how the sample controlled the subject's responding in general and his/her responding by exclusion in particular. We predicted that, given a choice between an undefined comparison stimulus (C_Y) and another defined stimulus (C_1 or C_2), responding by exclusion would be more likely as the duration of the undefined sample M_X moved away from the duration of the defined sample (i.e., the sample previously related to the defined comparison stimulus, M_{200} if C_1, or M_{600} if C_2). In other words, as the subject perceives the sample as less similar to the defined stimulus related to the alternative, defined comparison, the more likely the subject will be to respond by exclusion and choose the undefined comparison C_Y.

METHODS

Subjects

Ten university students, 18 to 26 years of age, participated in the experiment. They were naive with regard to experimental procedures related to temporal discrimination and exclusion. Data collection began after the ethics committee of the Federal University of São Carlos approved the study (protocol no. CAAE 0022.0.135.000-10) and the subjects signed an informed consent form.

Experimental settings and materials

The experimental sessions were conducted in a quiet room of the Laboratory of Studies of Human Behavior at the Federal University of São Carlos. The acoustically isolated room was equipped with one table, two chairs, a portable computer running the ProgMTS software (Marcicano, Carmo, & Prado, 2011), a mouse, two speakers, and a video camera.

Stimuli

The experimental task consisted of using the computer mouse to choose visual stimuli on the computer screen after the presentation of the acoustic sample. The sample stimuli were 500-Hz, 70-dB tones that varied in duration. The tones were presented through the speakers connected to the computer. The defined comparison stimuli were four circles with different colors: red (R), green (G), blue (B), and yellow (Y). The undefined comparison stimuli were six geometric, black-and-white shapes, henceforth designated by U_1, U_2, …,U_6. All comparison stimuli were 3.0 cm wide by 3.0 cm tall.

General procedure

The subjects first learned two temporal discriminations. In the first discrimination, the subject heard a 200 or 600 ms tone at the beginning of each trial. Then, two circles, Rand G, appeared in a random location on the computer screen. The choice of R was correct after the 200 ms tone and the choice of G was correct after the 600 ms tone (i.e., M_{200}→RandM$_{600}$→G). In the second discrimination, the procedure was similar, except that the subject heard a 600 or 1800 ms tone and then chose between two other circles, Band Y, with the correct mapping beingM$_{600}$→Band M_{1800}→Y. Importantly, the longer sample in the first discrimination was the shortest sample in the second discrimination. This task has been used with pigeons to study the effect of the learning context on temporal discriminations (see Machado, Malheiro, & Erlhagen, 2009; Oliveira &Machado, 2007; Machado & Pata, 2005), the interaction among generalization gradients (Vieira de Castro& Machado, 2010, 2012), and the emergence of new conditional relations (Huziwara, Velasco, Tomanari, de Souza, & Machado, 2012).

The study was divided into four training phases and one testing phase. Table 1 shows the structure of each training phase. To facilitate learning, in training Phases 1 and 2 the comparison stimuli were introduced gradually. Hence, during the first four trials of Phase 1, only the M_{200} stimulus was presented as a sample, and only the R circle (i.e., the correct response) was presented as the comparison. In the following eight trials, only the M_{200} stimulus continued to be presented as a sample, but now two comparison circles appeared. In the next four trials, only the M_{600}sample was presented, and only the G circle (the correct response) appeared on the computer screen. In the following eight trials, the two circles appeared. Lastly, in the following 36 trials, the two samples were presented 18 times each, in a random order, followed by the two comparisons, Rand G. Each correct response was followed by an oral expression of social approval presented by the computer (e.g., “Very good,” “Great,” “OK,” “Excellent,” or “Congratulations”).Incorrect responses did not have any programmed consequences. A 1-s interval separated consecutive trials.

—Table 1—

At the beginning of Phase 1, the experimenter gave the subjects the following instructions: “You will perform a task in which you have to use the mouse to click on a circle or on another stimulus when they appear on the screen after the beep.”

In Phase 2, all of the procedural details remained the same as in Phase 1, except that the samples and comparisons belonged to the second discrimination ($M_{600}\rightarrow$ Band $M_{1800}\rightarrow$ Y). In Phase 3, the four types of trials occurred in the same session in a random order. In addition, a criterion was introduced to advance to the next phase. If, for any of the four trial types, the subject made more incorrect than correct responses, then Phase 3 was repeated.

The goal of Phase 4 was to adapt the subjects to the absence of reinforcement that would be present during the subsequent test trials. To this end, the subjects received the following instructions: “You will have to perform the same task as before. However, not all of your responses will be praised.” For each sample, there were nine reinforceable trials (in case of correct choice) and 12 unreinforced trials (even if choice was correct). Phase 4 continued until the subjects made at most one error in each bisection task. Each subject required from 1to3 training sessions. Each session lasted 15 min, occurred generally once per day, and was video recorded.

Phase 4 was followed by the test phase that had the goal of determining how the subjects chose between a defined comparison stimulus (R, G, B, or Y) and an undefined comparison stimulus (U_1 to U_6) conditional on sample stimuli with different durations. The instructions were as follows: "You will continue to perform the same task of clicking one of the figures that appears on the screen after the beep."

Table 2 shows the structure of each session. The trials were of two sorts, baseline trials and test trials. The baseline trials were similar to the previous training trials and allowed us to monitor the accuracy of the base discriminations. The test trials included six samples varying in duration, three with the previously trained durations (i.e., 200, 600,and 1800 ms)and three with new durations (i.e., 350, 1040,and 7200 ms). Of the new durations, the first two corresponded to the geometric means of the pairs of trained durations (350 is approximately the geometric mean of 200 and 600, and 1040 is approximately the geometric mean of 600 and 1800); the 7200-ms test duration was substantially greater than the longest duration used during training. Each test sample was followed by two comparisons, an undefined comparison (shape) that remained the same for each sample, and one of the four defined comparisons (colors; see last column of Table 2). Each of the three previously trained durations was presented four times, one for each defined comparison, and each of the three new durations was presented 12 times, four times for each defined comparison. All test trials were performed in extinction, even those that used previously trained samples. In each testing session, the baseline and test trials occurred in random order. Five testing sessions were conducted, generally once per day, and lasted approximately 15 min.

At the end of the study, in a brief interview, the subjects responded to three questions: (1) “Was there any criterion to click on one or the other circle? If yes, what was it?”; (2) “What was your choice criterion when there was a colored circle and a figure that you had not seen before?”; and (3) “Did you count to measure time? If so, how did you do it?” The responses were recorded on video and analyzed subsequently. Finally, as a reward for participating in the study, the students received a coupon exchangeable for photocopies at the university store; the value of the coupon was the same for all subjects.

RESULTS

The average percentage of correct responses in training Phases 1, 2, and 3 was 89%. In Phase 4, with intermittent reinforcement of correct responses, the subjects needed an average of 2.3 sessions to meet the criterion, and the average percentage of correct responses equaled 97%. During the five testing sessions, correct responses on baseline trials equaled 80%. Although lower than in the training phases, the accuracy of the base discriminations remained high.

—Figure 1—

Figure 1shows the average performance on the exclusion tests. Consider the curve that corresponds to the R comparison (filled circles).Consistent with our hypothesis, the choice of the undefined stimulus was more likely as the duration of the sample moved away from 200 ms, the duration associated with R during training. The results for the G (empty circles) and B (filled triangles) comparisons, both associated with the 600 ms sample during training, revealed the same effect: The preference for the undefined stimulus was more likely as the duration of the sample moved away from 600 ms. Finally, the results for the Y comparison (empty triangles) also revealed an increasing preference for the undefined stimulus as the duration of the sample moved away from 1800 ms, although this increase was only substantial for shorter samples. In summary, the results of the four tests were consistent with the hypothesis that the more the test sample differed from the sample paired during training with the comparison present in the choice set(i.e., the more undefined the sample was with respect to the defined comparison stimulus), the more likely was the choice of the undefined comparison.

To evaluate the statistical significance of the results, we performed four analyses of variance (ANOVAs), one for each defined comparison stimulus, with duration as the repeated measure. All of the analyses revealed a significant effect of the duration of the sample on the preference for the undefined stimulus ($F_{5,45}>7.7$, $p<.01$). In this context, it is of particular interest to compare the two curves for the 600-mssample. Visual inspection of Figure 1 suggests that the curves are similar in shape but not in location because the B curve is shifted to the right compared with the G curve. A repeated-measures ANOVA with sample duration (6 levels) and comparison stimulus (2 levels) as factors revealed a significant interaction ($F_{5,45}=5.34$, $p<.01$).As described below, closer analysis of the G and B curves revealed two distinct groups of subjects, one that produced very similar, overlapping G and B curves, and another that produced a B curve clearly shifted to the right of the G curve.

To better understand the nature of individual differences, Figures 2 and 3 show the data from all subjects. Figure 2 shows the results of the four subjects that produced overlapping G and B curves. In the tests with these comparisons, the proportion of choices of the undefined stimulus followed a U-shaped curve, with a minimum at or close to 600 ms. In the tests with the R comparison, the four subjects displayed a systematically increasing curve. In the tests with the Y comparison, the differences among subjects occurred to a small extent at the 1800-ms sample and to a larger extent at the 7200-ms sample.

—Figure 2—

Figure 3 shows the results from the subjects that produced non-overlapping G and B curves. In the tests with these comparisons, five subjects (with the exception of S3 in the figure) produced a B curve to the right of the G curve. In the tests with the R comparison, all subjects showed an increasing curve, with a minimum at 200 ms. Finally, in the tests with the Y comparison, differences were again detected at the 1800 and 7200-ms samples.

—Figure 3—

In summary, with some exceptions, the average results presented in Figure 1 represent well the individual results. All subjects showed increasing curves in the tests with the R comparison. All subjects showed U- or V-shaped curves in the tests with the G comparison. Eight of the 10 subjects showed V- or U-shaped curves with the B comparison. These findings are consistent with our hypothesis. The data with the Y comparison were slightly more variable. Half of the subjects showed V-shaped curves with higher values at the short samples, minimal values at 1800 ms, and higher values at the longest duration of 7200 ms, a choice pattern clearly consistent with the hypothesis. However, the other five subjects showed curves with a minimum at 7200-ms samples, or with a minimum at1800-mssamples but with an equally low value at the 7200-ms samples, a result inconsistent with our hypothesis.

Verbal responses

The aforementioned results raise two questions. First, why was the B curve shifted to the right compared with the G curve for half of the subjects? Second, why did the Y curve show higher variability compared with the other curves, especially the R and G curves? Answers to these questions might be gleaned from the interview conducted at the end of the exclusion tests. In the Appendix, we include the complete answers of all subjects to the three questions asked by the experimenter. Here, we consider only the answers that refer to the choice of the comparison stimuli.

All subjects explicitly stated a rule for responding by exclusion that was consistent with our initial hypothesis (i.e., the choice of the undefined stimulus was more likely when the duration of the sample moved away from the duration associated with the colored stimulus). In what follows we illustrate this finding for some of the subjects:

> [S10]: "If the beep was similar to the beep presented with the color, I clicked on the color. When I did not know how much time the beep had, because sometimes I had the sensation that I had never heard that one before, I clicked on the geometric figure."

> [S8]: "I saw if the sound was equal to or different from the color. If it was equal, I chose the color. If the size of the sound was not associated with that color, I chose the figure."

[S9]: “If the time was different from the one that I remembered at the beginning when I clicked on the color, then I chose the little figure. If the time corresponded to a color, I clicked on the color.”

[S3]: “Sometimes the sound was very similar to the one played before, and I clicked on the color. Then sometimes the sound was very different from the one that I heard before. When it was different from the one that I had listened to, I clicked on the figure, but I did not know if it was correct. I just thought that it was not the color so I clicked on the other one.”

[S5]: “I saw if the time that appeared corresponded to the time that I had associated with the color before. If it was not similar to the one that I had associated, I clicked on that figure that had never appeared before.”

[S6]: “If the time was the one that I remembered, I clicked on the color. There was the short group that was red and green and the longer one that was blue and yellow. If the sound was big, for example, and had green and the figure, I clicked on the figure, got it? The most difficult part was when it had yellow and the sound was big.”

Some of the subjects stated that the B comparison was associated with a longer duration than the G comparison. Following are some examples:

[S4]: “The times were increasing. First was the red, then the green was longer, then the blue was a little longer, and then the yellow.”

[S5]: “It was more by the time. If I saw that the sound was short, I chose the red. If it was long, I went for the yellow. The green and blue confused me a bit. Sometimes I found that the green was shorter, sometimes not. But in general the green was smaller.”

[S6]: “It was more the duration of the sound that mattered to me, and I made the choice. It was like the sound was growing. Started with red, growing to green, then blue, and then yellow.”

[S7]: “I thought that the red was a little bit fast, the green a little bit longer than the red, and then came the blue and then the yellow. Then later I compared the two colors that appeared. If it was very short, it was red. If it was short, it was green. If it was a little longer, it was blue. If it was very long, it was the yellow.”

[S9]: “The red was fast, the green a little less fast, the blue even less fast, and the yellow was longer. In fact, the red was linked to the green and was shorter, and the blue was linked with the yellow and was longer.”

[S10]: “I ‘gave’ a time to each color. For example, the red had 1 second, the yellow had between 3 and 4 seconds, and the green and blue were in between. I could not distinguish between the green and blue, but if it was a little bit longer, it was blue because it was together with the yellowat the beginning."

The previous statements suggest that some of the subjects divided the comparisons into two classes according to the (subjective) duration of the corresponding sample. One class contained the R and G stimuli and was associated with shorter samples, and another class contained the B and Y stimuli and was associated with longer samples. Curiously, several subjects seemed to have perceived the sample that corresponded to B as a bit longer than the sample that corresponded to G, a result that can explain the shift of the B curve to the right of the G curve.

However, there were several discrepancies between the verbal answers and the results of the choice tests, and they suggest that the Y comparison could have been reallocated to the longest sample. For example, subject S2 stated in the interview that "if the size of the sound was closer, more like the color, I clicked on the color. If it was very different from the color, I clicked on the figure."If we take into account that this subject discriminated the base durations reasonably well during training (overall percent correct =77%), we would predict that in the exclusion tests the subject would prefer the Y stimulus after the 1800-ms training sample, but would prefer the undefined stimulus after the substantially longer 7200-ms untrained sample. However, this result did not occur (see Figure 2). Thus, the verbal report of this subject is consistent with the R, G, and B curves but not Y curve, unless the Y stimulus is reallocated to the longest sample.

The cases of subjects S4 and S5 are similar. Although they both had good performance during training, with more than 79% correct choices after the 1800 ms sample, and verbally reported that they chose the undefined stimulus when "the sound associated with the color did not show up" (S4) or when the "[duration] did not resemble the one that I had associated” (S5), both always chose the undefined stimulus at 1800 ms and the Y stimulus at 7200 ms. In summary, for some subjects, the Y stimulus associated with the longest (1800-ms)sample during training may have been reallocated to the longest(7200-ms) sample during testing.

DISCUSSION

The goal of the present study was to investigate responding by exclusion using simple sample stimuli, that is, stimuli that differed along a single dimension. To that end, we studied performance in conditional discriminations, matching-to-sample tasks that included samples differing only in duration. After the subjects learned to associate specific sample durations (tones) with specific comparison stimuli (colored circles), they were tested with novel sample durations followed by novel comparison sets. The latter included always one of the comparisons defined during training (the R, G, B, orY circles) and a new, undefined comparison (the U_1 to U_6 geometric shapes).

The average results showed that the probability of choosing the undefined stimulus when presented with the R comparison, previously associated with the 200-ms sample, was minimal at 200msand increased with the sample duration. When the undefined stimulus was presented with the Y comparison, previously associated with the 1800-ms sample, the opposite trend was observed, that is, preference for the undefined stimulus was strongest at the shortest durations and decreased, on average, at the longest1800-and 7200-ms durations. When the undefined stimulus was presented with the G or B comparisons, preference for the undefined stimulus followed a U- or V-shaped curve (i.e., the probability of choosing the undefined stimulus was higher at the shortest and longest durations and lower at durations close to 600 ms, the sample previously associated with these comparisons).

These results reveal the accuracy of the trained discriminations because the R comparison was conditionally trained with the shortest sample, and the Y comparison was conditionally trained with the longest sample. With respect to the G and B comparisons, the similarity of the curve profiles indicate that the subjects tended not to chose the undefined stimulus after the samples paired with G and B during training, and to choose it when the test sample were judged as sufficiently different from the samples paired with G and B during training. These results replicate the findings obtained with concrete stimuli that vary along multiple dimensions (e.g., Bates et al., 1979; Carey & Bartlett, 1978; Costa et al., 2001; Dixon, 1977; McIlvane et al., 1992; Stromer & Osborne, 1982; Wilkinson et al., 2009).

The procedure used in the present study differed from the procedures used in previous studies also in the composition of the test trials. Typically, responding by exclusion is evaluated on trials in which the undefined comparison C_Y is presented as a choice alternative next to a defined stimulus, which defined stimulus may be the S^- for that sample because it was paired with a different sample during training). With this arrangement of comparisons (C_Y, C_i), the subject excludes the previously defined relations and thereby chooses the undefined stimulus. In contrast, in the present study, if for each sample (e.g., 350 ms) we order the four comparisons, starting with the comparison associated with the fewest choices of the undefined stimulus and ending with the comparison associated with the most choices of the undefined stimulus (e.g., in Figure 1, for the 350 ms duration, the order would be R, G, B, and Y),then the order is consistent with the order predicted by the temporal generalizations gradients. In other words, the average ordering is consistent with the idea that the choice of the undefined comparison C_Y varied inversely with the difference $|M_x-M_i|$, the difference between the duration of the test sample, M_x, and the duration of the sample associated with the defined comparison, M_i.

The foregoing conclusions based on choice proportions generally were consistent with the subjects' verbal reports. Most subjects (S2, S3, S4, S5, S8,and S9) said that when faced with a sample never seen before, their choice was for the "different one"(i.e., the undefined stimulus). Some subjects (S1, S6, and S7) chose the undefined stimulus when the defined comparison in the set was considered "shorter" (R, G) and the sample was considered longer. Similarly, they clicked the undefined stimulus when the sample was considered "short" and the defined comparison was B or Y. Notice that even in this case, the subjects still chose the undefined comparison by excluding the defined comparison because the latter was not associated during training with the test sample.

Responding by exclusion, that is, choosing an undefined comparison following an undefined sample can take place under two distinct conditions. The subject may reject the defined comparison or the subject may select the undefined comparison. These two conditions are accompanied by the same observable behavior but they imply two stimulus classes, one comprising the defined comparisons and the other comprising the undefined comparisons (McIlvane, Wilkinson, & de Souza, 2000; Wilkinson & McIlvane, 1997). Hence, one of the limits of responding by exclusion stems from the specific features of the comparisons. To illustrate, Costa (2008) and Costa, de Rose, and de Souza (2010) showed that the presence of suffixes in undefined verbal samples (e.g., a diminutive, as in "mopadinho" instead of "mopadi") may reduce responding by exclusion. Presumably because that lexical cue had been related during the subject's history of verbal behavior to objects of small size, that cue may have competed with the usual stimulus control sources of responding by exclusion. More generally, then, to the extent that one or more properties of the sample and comparison stimuli have been related to other stimulus classes (as diminutives have been related to the class of small objects), responding by exclusion is less likely to occur.

The present study also contributed to studies of exclusion by using temporal stimuli as samples, that is, samples that varied along a single dimension. Its results show that responding by exclusion may be controlled by a single stimulus property (i.e., duration). More generally, studies with different stimulus modalities can improve our knowledge about the responding by exclusion. The present study found preferential choice of the new stimulus when the sample also was new, a result consistent with other studies using concrete stimuli. However, the novelty of an undefined stimulus depended on the degree of generalization related to the previously defined stimuli. Although 350-ms samples were not in the trained set, the majority of subjects responded to it as they did to 200-ms samples, preferring the R comparison when present (i.e., the one that was trained as the correct stimulus in short trials) when it was present. The results of the present study permit accurately evaluating and establishing the limits of responding by exclusion that have been indicated by previous studies (e.g. Costa, 2008; Costa, de Rose, and de Souza, 2010).

REFERENCES

Aust, U., Range, F., Steurer, M., & Huber, L. (2008).Inferential reasoning by exclusion in pigeons, dogs, and humans.*Animal Cognition*, *11*, 587-597.

Bates, E., Benigni, L., Bretherton, I., Camaioni, I., &Volterra, V. (1979).*The emergence of symbols: cognition and communication in infancy*. New York: Academic Press.

Beran, M. J., & Washburn, D. A. (2002). Chimpanzee responding during matching to sample: control by exclusion. *Journal of the Experimental Analysis of Behavior*, *78*, 497-508.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development*, *15*, 17-29.

Clement, T. S., & Zentall, T. R. (2000). Development of a single-code/default coding strategy in pigeons.*Psychological Science*, *11*, 261-264.

Clement, T. S., & Zentall, T. R. (2003). Choice based on exclusion in pigeons. *PsychonomicBulletin&Review*, *10*, 959-964.

Costa, A. R. A. (2008). *Investigação sobre exclusão na aquisição de relações condicionais auditivo-visuais por crianças com desenvolvimento típico e por pessoas com atraso na aquisição de vocabulário.* Post doctoral fellow ship report submitted to FAPESP, Universidade Federal de São Carlos.

Costa, A. R. A., de Rose, J. C., & de Souza, D. G. (2010). Interferência de variáveis de contexto em sondas de exclusão com substantivos e verbos novos. *Acta Comportamentalia*, *18*, 35-54.

Costa, A. R. A., Wilkinson, K. M., McIlvane, W. J., & de Souza, D. G. (2001). Emergent word-object mapping by children: further studies using the blank comparison technique. *Psychological Record*, *51*, 343-355.

Cumming, W. W., & Berryman, R. (1961). Some data on matching behavior in the pigeon.*Journal of the Experimental Analysis of Behavior, 4*, 281-284.

Dixon, L. S. (1977). The nature of control by spoken words over visual stimulus selection.*Journal of the Experimental Analysis of Behavior*, *27*, 433-442.

Dixon, M. H., Dixon, L. S., &Spradlin, J. E. (1983). Analysis of individual differences of stimulus control among developmentally disabled children. In K. D. Gadow, & I. Bialer (Eds.), *Advances in learning and behavioral disabilities*, vol. 2 (pp. 85-110). Greenwich, C.T.: JAI Press.

Domeniconi, C., Costa, A. R. A., de Souza, D. G., & de Rose, J.C. (2007). Responder por exclusão em crianças de 2 a 3 anos em uma situação de brincadeira. *Psicologia Reflexão e Crítica*, *20*, 342-350.

Ferrari, C., de Rose, J. C. C., & McIlvane, W. J. (1993). Exclusion vs. selection training of auditory-visual conditional relations.*Journal of Experimental Child Psychology, 56*, 49-63.

Floor, P., &Akhtar, N. (2006). Can 18-month-old infants learn words by listening in on conversations? *Infancy*, *93*, 327-339.

Halberda, J. (2003). The development of a word-learning strategy.*Cognition,* 87, B23-B34.

Herman, L. M., Richards, D. G. and Wolz, J. P. (1984). Comprehension of sentences by bottlenosed dolphins.Cognition 16: 129-219.

Huziwara, E. M., Velasco, S. M., Tomanari, G. Y., de Souza, D.G., & Machado, A. D. (2012). Emergent relations in pigeons following training with temporal samples.Learning & Behavior, *40*, 192-204.

Kaminski, J., Call, J., & Fischer, J. (2004). Word learning in a domestic dog: evidence for "fast mapping." *Science*, *304*, 1682-1683.

Kastak, C. R., &Schusterman, R. J. (2002). Sea lions and equivalence: expanding classes by exclusion. *Journal of the Experimental Analysis of Behavior*, *78*, 449-465.

Machado, A., &Pata, P. (2005) Testing the scalar expectancy theory (SET) and the learning-to-time sample (LeT) in a double bisection task. *Learning &Behaviour*, *33*, 111-122.

Machado, A., Malheiro, M.T. &Erlhagen, W. (2009). Learning to time: a perspective. *Journal of the Experimental Analysis of Behavior*, 92, 423-458.

Marcicano, D. C., Carmo, J. S., & Prado, P. S. T. (2011).Software ProgMTS: Possibilidades de delineamento e condução de programas de ensinoemAnálise Experimental do Comportamento [programa de computador]. In proceedings of the 41st Annual Meeting of the Brazilian Society of Psychology: Training and Production of Knowledge in Psychology, Bethlehem.

McIlvane, W. J., Kledaras, J. B., Lowry, M. J., & Stoddard, L. T. (1992).Studies of exclusion in individuals with severe mental retardation. *Research in Developmental Disabilities*, *13*, 509-532.

McIlvane, W. J., Kledaras, J. B., Munson, L. C., King, K. A., de Rose, J. C., & Stoddard, L. T. (1987). Controlling relations in conditional discrimination and matching by exclusion. *Journal of the Experimental Analysis of Behavior*, *48*, 187-208.

McIlvane, W. J., Munson, L. C., & Stoddard, L. T. (1988).Some observations on control by spoken words in children's conditional discrimination and matching by exclusion.*Journal of Experimental Child Psychology*, *45*, 472-495.

McIlvane, W. J., Wilkinson, K. M., & de Souza, D. G. (2000). As origens da exclusão. *TemasemPsicologia*, *8*, 195-203.

McIlvane, W.J., & Stoddard, L.T. (1981). Acquisition of matching-to-sample performances in severe mental retardation: learning by exclusion. *Journal of Mental Deficiency Research*, *25*, 33-48.

Oliveira, L., & Machado, A. (2007). The effect of sample duration and cue on a double temporal discrimination.*Learning and Motivation*, *39*, 71-94.

Oshiro, C. K. B., de Souza, D. G., & Costa, A. R. A. (2006).Responder porexclusão a partir de uma linha de base de discriminaçõescondicionaisvisuais. *Revista Brasileira de Análise do Comportamento*, *2*, 251-276.

Pruden, S. M., Hirsh-Pasek, K., Golinkoff, R. M., & Hennon E. A. (2006). The birth of words: ten-month-olds learn words through perceptual salience. *Child Development*, 77, 266-280.

Schloegl C, Bugnyar T, Aust U. Exclusion performances in non-human animals: from pigeons to chimpanzees and back again. In: Blaisdell A, Huber L, Watanabe S, Young A, Yamazaki Y, editors. Rational animals, irrational humans. Tokyo: Keio University Press; 2009. pp. 217–234.

Schusterman, R. J., & Krieger, K. (1984). California sea lions are capable of semantic comprehension. *Psychological Record*, *34*, 3-23.

Stromer, R. & Osborne, J. G. (1982). Control of adolescents arbitrary matching-to sample by positive and negative stimulus relations.Journal of the Experimental Analysis of Behavior, 37, 329-348.

Stromer, R. & Osborne, J. G. (1982). Control of adolescents arbitrary matching-to-sample by positive and negative stimulus relations. *Journal of the Experimental Analysis of Behavior, 37*, 329-348.

Stromer, R. (1986). Control by exclusion in arbitrary matching-to-sample. *Analysis and Intervention in Developmental Disabilities, 6*, 59-72.

Stromer, R. (1989). Symmetry of control by exclusion in humans' arbitrary matching-to-sample.*Psychological Reports, 64*, 915-922.

Vieira de Castro, A.C. & Machado, A. (2010) Prospective timing in pigeons: Isolating temporal perception in the time-left procedure. *Behavioural Processes, 84*, 490-499.

Vieira de Castro, A.C., & Machado, A. (2012) The interaction of temporal generalization gradients predicts the context effect. *Journal of the Experimental Analysis of Behavior, 97*, 263-279.

Wilkinson, K. M., de Souza, D., McIlvane, W. J. (2009). As Origens da Exclusão. *Temas em Psicologia da SBP, 8*, 195-203.

Wilkinson, K. M., Rosenquist, C., &McIlvane, W. J. (2009). Exclusion learning and emergent symbolic category formation in individuals with severe language impairments and intellectual disabilities.*Psychological Record, 59*, 187-206.

AUTHOR NOTES

Address correspondence to Camila Domeniconi, Programa de Pós-Graduação em Psicologia, Universidade Federal de São Carlos, Rod. Washington Luís, km 253. Caixa Postal 676, 13565-905, São Carlos, SP; email: domeniconicamila@gmail.com

The first author was supported by a master´s degree fellowship by the Ministry of Education (CAPES). Camila Domeniconi had a post doctoral fellowship by Foundation for Research Support in the State of São Paulo (FAPESP, 2009/18479-5). Armando Machado was supported a grant from the Portuguese Foundation for Science and Technology (FCT).

Camila Domeniconi is currently affiliated to Instituto Nacional de Ciência e Tecnologia sobre Comportamento, Cognição e Ensino, supported by FAPESP (Grant # 08/57705-8) and CNPq (Grant # 573972/2008-7), that provided support for preparation of this manuscript.

APPENDIX

At the end of the exclusion tests, the experimenter asked the subjects the following three questions:

1) "Was there any criterion to click on one or the other circle? If yes, what was it?";

2) "What was your choice criterion when there was a colored circle and a figure that you had not seen before?", and

3) "Did you count to measure time? If so, how did you do it?".

The answers of each subject are presented below.

Subject S1

1. "At the beginning, I started clicking on the red, and then I continued clicking on the same color because the tone that was played was equal. Then came the green and the tone was played longer. When I did this, when the tone played one way, I clicked on one, and if it played another way, I clicked on the other."
2. "Then, when this started I did not understand. I first clicked on the figure that I never saw before, and nothing happened. Then I kind of determined that when the red or green appeared, those in the beginning were shorter, and a long sound was played, I clicked on the figure because red and green were short. When the sound was a little short and had blue and yellow, I also chose the figure."
3. "No, I did not count the time, I saw that it was short or long and clicked."

Subject S2

1. "I did this, not knowing if it was right, but subdivided the sound and compared it with one color or another in relation to the size of the sound. Buteach color did

not always have the specific size. Like, I assigned that red was 0.5 to 1.0. The blue was 1.5 to 2.0. They varied in this interval that I determined, got it?”

2. “I saw it like this: If the size of the sound was close, it was more similar to a color, and I clicked on the color. If it was very different, I clicked on the figure.”
3. “I counted. The sound itself had oscillations. I counted up to 14 oscillations as the longest, then I was seeing how many and compared them.”

Subject S3

1. “There were sounds that were very small and others very big. Then there were the ones in the middle. The ones in the middle, I always got confused, but I paid attention to how long the sound was played and chose.”
2. “Sometimes the sound was very similar to the one played before, and I clicked on the color. Then sometimes the sound was very different from the one that I heard before. When it was different from the one that I had listened to, I clicked on the figure, but I did not know if it was correct. I just thought that it was not the color so I clicked on the other one.”
3. “Ah, it was fast, no time to count, no. Was I supposed to count?”

Subject S4

1. “The times were increasing. First was the red, then the green was longer, then the blue was a little longer, and then the yellow. In fact, I saw how long the sound played.”
2. “When the sound that I had associated with the color did not appear, I clicked on the other one, the one that did not have color and had a figure inside.”

3. “I was thinking 1, 2, 3,and counted in my head. But there were some that I could not, no way. When I tried, it was already done. I only counted the longer ones.”

Subject S5

1. “I was going more by the time. If I saw that the sound was short, I chose the red. If it was long, I went for the yellow. The green and blue got me a little bit confused. Sometimes I thought that the green was shorter, sometimes not. But in general the green was smaller.”
2. “I saw if the time that appeared corresponded to the time I had associated before with the color. If it was not similar to the one I had associated, I clicked on that figure that had never appeared before.”
3. “Ah, yes, I counted. In my head, when the tone started I started1, 2, 3, and so on.”

Subject S6

1. “It was more the duration of the sound that mattered to me, and I made the choice. It was like the sound was growing. Started with red, growing to green, then blue, and then yellow.”
2. “If the time was the one that I remembered, I clicked on the color. There was the short group that was red and green and the longer one that was blue and yellow. If the sound was big, for example, and had green and the figure, I clicked on the figure, got it? The most difficult part was when it had yellow and the sound was big.”
3. I counted nothing. I only felt if the sound was fast or slow.”

Subject S7

1. "I thought that the red was a little bit fast, the green a little bit longer than the red, and then came the blue and then the yellow. Then later I compared the two colors that appeared. If it was very short, it was red. If it was short, it was green. If it was a little longer, it was blue. If it was very long, it was the yellow."
2. "At the beginning, I started realizing that when the sound was very long, the yellow and a figure like a star appeared. Then I always clicked on the star because the yellow was not that long. Then after this no longer worked, I realized that there were shorter and longer sounds. If the short sound played and green or red appeared, I pushed the color because red and green were a small pair. When green or red appeared and the sound was longer, I clicked the figure because I thought that it [the tone] was not of the small pair green/red but rather of the blue/yellow."
3. "I saw that sometimes when the sound played, it had some little parts higher, like little jumps. Then I counted that. Sometimes I could not do that, but sometimes I could."

Subject S8

1. "By my logic, it was the duration of the sound that determined if I chose one or the other. If the sound was long, I pushed blue or yellow, if it was faster, then it was green or red. And between the red and green, the red was fastest, and between the blue and yellow, the yellow was slower."
2. "I saw if the sound was equal to or different from the color. If it was equal, I chose the color. If the size of the sound was not associated with that color, I chose the figure."

3. I did not count because it was too fast; one couldn't count."

Subject S9

1. "The red was fast, the green a little less fast, the blue even less fast, and the yellow was longer. In fact, the red was linked to the green and was shorter, and the blue was linked with the yellow and was longer."
2. "If the time was different from the one that I remembered at the beginning when I clicked on the color, then I chose the little figure. If the time corresponded to a color, I clicked on the color."
3. "At the beginning, I counted in my head, but after too many sounds began to appear, I couldn't count anymore."

Subject S10

1. "I 'gave' a time to each color. For example, the red had 1 second, the yellow had between 3 and 4 seconds, and the green and blue were in between. I could not distinguish between the green and blue, but if it was a little bit longer, it was blue because it was together with the yellow at the beginning."
2. "If the beep was similar to the beep presented with the color, I clicked on the color. When I did not know how much time the beep had, because sometimes I had the sensation that I had never heard that one before, I clicked on the geometric figure."
3. "I tried to count in my head how many seconds elapsed. It was not exactly the correct seconds, but I could count."

Table 1. Number of trials, sample and comparison stimuli presented in each training phase. InPhase 4, the two trial numbers indicate the number of reinforceable(9) andnonreinforced (12) trials. Sample duration is in milliseconds (e.g., M_{200}=200 ms). The comparison stimuli were circles of different colors (R=red,G=green,B=blue, and Y=yellow).

Phase	Trials	Sample	Comparisons
1	4	M_{200}	R
	8	M_{200}	R + G
	4	M_{600}	G
	8	M_{600}	R + G
	36	M_{200} or M_{600}	R + G
2	4	M_{600}	B
	8	M_{600}	B + Y
	4	M_{1800}	Y
	8	M_{1800}	B + Y
	36	M_{600} or M_{1800}	B + Y
3	8	M_{200}	R + G
	8	M_{600}	R + G
	8	M_{600}	B + Y
	8	M_{1800}	B + Y
4	9+12	M_{200}	R + G
	9+12	M_{600}	R + G
	9+12	M_{600}	B + Y
	9+12	M_{1800}	B + Y

Table 2. Number of trials, sample and comparison stimuli presented in each session of the testing phase. On the exclusion test trials, U_1, U_2,…,U_6refer to the six undefined comparisons (black and white geometric shapes) and C_i, with i=R, G, B, or Y, refers to the four defined stimuli (circles of different colors). The frequency of each C_i was equal to the number of trials divided by four.

Baseline			Exclusion Test		
Trials	Samples	Comparisons	Trials	Samples	Comparisons
9	M_{200}	R + G	4	M_{200}	$C_i + U_1$
9	M_{600}	R + G	12	M_{350}	$C_i + U_2$
9	M_{600}	B + Y	4	M_{600}	$C_i + U_3$
9	M_{1800}	B + Y	12	M_{1040}	$C_i + U_4$
			4	M_{1800}	$C_i + U_5$
			12	M_{7200}	$C_i + U_6$

FIGURE CAPTIONS

Figure 1. The curves show the average proportion of choices of the undefined stimulus as a function of sample duration with the defined comparison stimulus (R, G, B and Y) as a parameter. The values in parenthesis show the sample durations associated during training with the respective comparisons. The x-axis is on a log scale.

Figure 2. The curves sow the average proportion of choices of the undefined stimulus as a function of the sample duration, with the defined comparative stimulus (R, G, B, or Y) as a parameter. The duration of the model is on a logarithmic scale.

Figure 3. The curves sow the average proportion of choices of the undefined stimulus as a function of the sample duration, with the defined comparative stimulus (R, G, B, or Y) as a parameter. The duration of the model is on a logarithmic scale.