Universidade do Minho
Escola de Psicologia

Carlos André Ribeiro Oliveira Pinto

**Learning and coding strategies in pigeons
(*Columba livia*)**

março de 2016

Universidade do Minho

Escola de Psicologia

Carlos André Ribeiro Oliveira Pinto

Learning and coding strategies in pigeons (*Columba livia*)

Tese de Doutoramento em Psicologia Básica

Trabalho efetuado sob a orientação do
Professor Doutor Armando Machado

março de 2016

STATEMENT OF INTEGRITY

I hereby declare having conducted my thesis with integrity. I confirm that I have not used plagiarism or any form of falsification of results in the process of the thesis elaboration. I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

University of Minho, March 29, 2016

Carlos André Ribeiro Oliveira Pinto

Signature:

Everything is vague to a degree you do not realize till you have tried to make it precise.

Bertrand Russell, *The Philosophy of Logical Atomism*

The way to deal with an impossible task was to chop it down into a number of merely very difficult tasks, and break each one of them into a group of horribly hard tasks, and each of them into tricky jobs, and each of them…

Terry Pratchett, *Truckers*

Para as minhas Marias.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Professor Armando Machado, for all the doors he opened to me, and for illuminating what lies beyond many of those doors.

LEARNING AND CODING STRATEGIES IN PIGEONS (*COLUMBA LIVIA*)

Abstract

In a series of five studies, we analyzed discrimination learning in pigeons by focusing on which stimuli the animals would attend to, and, based on those stimuli what response rules (or codes) would be created. We used a matching-to-sample task that featured three samples and two comparisons. One comparison was correct following one of the samples and the other comparison was correct following the remaining samples. This arrangement allowed the task to be solved in more than one way. One solution would consist in establishing one code per sample duration, totaling three codes. We were interested in assessing whether the pigeons would resort to a single-code / default strategy that would require only two codes: one code for the sample that was matched to an exclusive comparison, and another code to be applied to all other samples. When the task involved a temporal discrimination, the results suggested that both the sample and houselight (that signaled the intertrial interval) controlled responding, and that the pigeons did not create three sample-specific codes. However, the evidence was not conclusive in clarifying if the coding strategy employed was the single-code / default. On the other hand, when the task involved a non-temporal discrimination, the majority of birds did not show evidence of single-code / default. Taken together, our results suggest that the pigeons were flexible enough to adapt to the specificities of the task. Additionally, the conditions that trigger the adoption of a given coding strategy may not only depend on the characteristics of the task and the modality of the stimuli used, but can also vary substantially among individuals.

ESTRATÉGIAS DE APRENDIZAGEM E CODIFICAÇÃO EM POMBOS (*COLUMBA LIVIA*)

Resumo

Num conjunto de cinco estudos, analisámos a aprendizagem de pombos em tarefas discriminativas visando identificar a que estímulos o animal atenderia e, com base nesses estímulos, que regras de resposta (ou códigos) criaria. Usámos uma tarefa de correspondência à amostra com três amostras e duas comparações, em que uma comparação era correta após uma das amostras e a outra comparação era correta após as restantes amostras. Esta estrutura permitia que a tarefa fosse aprendida de mais que uma forma. Uma forma consistiria no estabelecimento de um código por amostra, num total de três códigos. Estávamos interessados em testar se os animais fariam uso de uma outra estratégia, a "single-code / default", que envolveria o estabelecimento de apenas dois códigos: um para a amostra que foi associada a uma comparação exclusiva, e um outro código que seria aplicado a todas as outras amostras. Quando a tarefa envolveu uma discriminação temporal, tanto a amostra como a luz de iluminação geral (que sinalizava o intervalo entre-ensaios) pareceram controlar as respostas, e os pombos não criaram três códigos específicos, um para cada amostra. Contudo, os dados não permitiram clarificar de forma conclusiva se a estratégia de codificação em uso foi a "single-code / default". Por outro lado, quando a tarefa envolveu uma discriminação não temporal, a maioria dos pombos não mostrou evidência de uso de "single-code / default". Em conjunto, os nossos dados sugerem que os pombos mostraram flexibilidade para adaptarem-se às particularidades da tarefa. Adicionalmente, as condições para a adopção de uma estratégia podem não só depender das características da tarefa ou da modalidade dos estímulos, como também podem variar significativamente entre indivíduos.

TABLE OF CONTENTS

STUDY IV - CODING IN PIGEONS: FURTHER TESTS OF SAMPLE-COMPARISON MAPPING

STUDY V - CODING IN PIGEONS: BOUNDARY CONDITIONS OF THE SINGLE-CODE / DEFAULT STRATEGY IN A VISUAL DISCRIMINATION TASK

ABBREVIATIONS, ACRONYMS AND SYMBOLS

α – Training performance (Study III)

ABET – Animal Behavior Environment Test

ANOVA – Analysis of Variance

C – Comparison stimulus

CI – Confidence Interval

h – height

IEE – Industrial Electronics Engineers

ITI – Intertrial Interval

l – length

LVE – Lehigh Valley Electronics

p – Probability of using houselight offset as time marker

PVC – Polyvinyl chloride

S – Sample stimulus

SEM – Standard Error of the Mean

w – width

Figure 8. Schematic of the stimuli presented during a trial in Training (top panel) and Retention Tests (bottom panel). Each line is raised whenever its corresponding stimulus was turned on.

Figure 9. Mean (with SEM) percent correct to each sample duration as a function of retention interval duration. The panels on the left refer to Retention Test I and the panels on the right refer to Retention Test II. The panels on the top refer to the group of pigeons that first learned the task with two samples and then learned the task with three samples. The panels on the bottom refer to the group of pigeons that learned the tasks in reverse order.

Figure 10. Mean (with SEM) percent of choices to the "long" key (associated with 6-s and 18-s samples) as a function of retention interval on the Lit ITI, No-sample Retention Test (empty circles) and on the Dark ITI, No-sample Retention Test (filled circles).

Figure 11. Mean percent of choices to the "long" key (associated with 6-s and 18-s samples) as a function of time since houselight offset. Each line refers to a sample duration. The filled circles represent the collapsed data from Retention Test I and II and the empty circles represent the data from the Lit ITI, No-sample Retention Test.

Figure 12. Mean (with SEM) proportion correct to each sample duration during testing, in training trials (left side) and in dark-ITI test trials (right side).

Figure 13. Proportion of choices to the "short" key in no-sample test trials (x axis) plotted against proportion of correct responses following 2-s samples in dark-ITI test trials (y axis). The line is the prediction of the model (with $\alpha = .94$). Each dot refers to the performance of one pigeon. The white dot identifies pigeon PG29.

Figure 14. Mean (with SEM) percent correct following each of the three sample durations as a function of retention interval. 2-s and 18-s samples (filled lines) shared the correct comparison.

INTRODUCTION

The environment in which any organism lives, despite its complexity and variability, has many regularities. Understanding these regularities is important because it allows making predictions based on present information. Therefore, the ability to identify regularities and translate them into response rules that allow an organism to behave adaptively is essential. The response rules an animal establishes are known as codes, and the process of learning them is known as coding. To study coding in a laboratory setting, a simplified environment where regularities can be controlled is used– an operant chamber. This instrument allows the presentation of different tasks, to study how an animal learns and adapts to them.

One of the most commonly-used tasks to study coding is the matching-to-sample task. To illustrate, picture a task that starts with the presentation of one of two stimuli (samples) that differ in duration, say, a light that stays on for 2 s or 10 s. Following this initial stimulus, a choice is given between two colored keys (comparisons), say, green and red. The green key is correct if the light was on for 2 s, and the red key is correct if the light was on for 10 s. To be able to choose the correct comparison, an animal needs to learn the aforementioned rules governing the task, and to behave according to them.

The preceding matching-to-sample task is symmetrical, in the sense that the number of samples equals the number of comparisons, and they are mapped one-to-one, (i.e., for each sample there is only one correct comparison, and each comparison is correct following only one of the samples). In this symmetrical task, is coding also symmetrical, that is, do both samples exert control over responding such that a response rule is established for each one of them (e.g., "if 2 s choose green", "if 10 s choose red")? Or is coding asymmetrical, with only one of the samples controlling choice and with response rules established based on that sample (e.g., "if 2 s choose green", "if not 2 s choose red")?

To address these questions, Grant and Spetch (1994) trained pigeons on two symmetrical matching tasks simultaneously. One task had 2-s and 10-s samples and green and red as comparisons, and the other task had 4.5-s and 22.5-s samples and vertical and horizontal lines as comparisons. Following training, the authors introduced trials with components from the two tasks intermixed (e.g. 2-s or 4.5-s samples

followed by a choice between one color and one line orientation). Performance on these mixed trials was above chance for all four sample durations, which suggests that pigeons had learned a specific response rule for each sample; coding was seemingly symmetrical. Other studies suggest that in asymmetrical tasks, where more than one sample is associated with each comparison, a common code is established for the samples that share the same comparison (e.g., Grant & Spetch, 1993; Urcuioli, Zentall, Jackson-Smith, & Steirn, 1989; Zentall et al., 1992; Zentall, Steirn, Sherburne, & Urcuioli, 1991).

One of the goals of the present dissertation was to study coding flexibility, by using a task that was a mixture of the symmetrical and asymmetrical tasks described above: three samples map onto two comparisons, with one comparison correct following one of the samples (one-to-one mapping), and the other comparison correct following the remaining two samples (many-to-one mapping). This task can be learned symmetrically, by establishing three codes, one for each sample, or asymmetrically, by using a strategy known as single-code / default: a specific code is established for the sample on one-one correspondence (single code), and a default code is established for the remaining samples on many-one correspondence. Since the asymmetrical coding involves fewer codes than the symmetrical counterpart, it is arguably the more efficient of the two strategies.

There is evidence suggesting that animals are able to adjust their coding strategies depending on the task, with evidence of single-code / default in tasks where the sample stimuli were the presence or absence of a stimulus (e.g., Colwill, 1984; Grant, 1991; Sherburne & Zentall, 1993; Weaver, Dorrance, & Zentall, 1999; Wilkie, 1978; Wilson & Boakes, 1985), or samples differed in the amount of training (Grant, 2006a; Grant & Blatz, 2004) or in salience (Grant, 2009b; Wixted & Gaitan, 2004), so we wanted to confirm whether pigeons would adopt the most economical of the two coding strategies in the 3-sample, 2-comparison task.

A method commonly used to study whether many-to-one tasks promote single-code / default coding consists in introducing a retention interval between sample offset and comparison onset after training the task. The rationale is that accurate performance requires the retention of information related to the sample (or to the comparison to be chosen) throughout the retention interval. The longer the retention interval, the more likely it is that the retained information is lost. If an animal is using a single-code / default strategy, when faced with a choice following the loss of information, it should

resort to the "default" response (since there is no information available regarding the sample for which the "single code" was established). Therefore, if an animal has adopted a single-code / default strategy, following long-enough retention intervals there should be a preponderance of choices to the "default" comparison, the comparison with the many-to-one mapping.

Gaitan and Wixted (2000) trained pigeons on two tasks where three samples, lasting 0, 2 and 10 s, mapped onto two colored comparisons: in one task, 0-s and 2-s samples shared a correct comparison (Experiment 2); on the other task, 0-s and 10-s samples shared a comparison (Experiment 3). The results of subsequent retention tests were consistent with single-code / default coding: In both experiments, there was a preference for the comparison with the many-to-one mapping, what would be the "default" comparison. However, this evidence was not definitive. The interval that separated trials was spent in darkness, so in trials with 0-s samples, the pigeons could have learned to choose one comparison following a period of darkness. Since the retention interval was also spent in darkness, the results of the retention tests could be due to the birds selecting the comparison that, in training, was correct immediately following a period of darkness. In both experiments, the 0-s sample was part of the many-to-one mapping, so a preference for the 0-s comparison and the predictions of the single-code / default strategy were confounded.

To assess whether the similitude between intertrial and retention intervals was of significance, Zentall, Klein and Singer (2004) replicated Gaitan and Wixted's (2000) Experiment 2, one condition with intertrial and retention intervals dark, and other condition with an illuminated intertrial interval. When the intervals were similar, a preference for the comparison associated with 0- and 2-s was found; when the intervals were different there was no difference in preference between comparisons. These results suggest that the evidence for single-code / default obtained by Gaitan and Wixted (2000) could in fact be an artifact introduced by the use of 0-s samples.

To avoid potential confounds with 0-s samples, Singer, Klein and Zentall (2006, Experiment 2) trained pigeons with three nonzero sample durations (2, 8 and 32 s). In their task, 2- and 32-s samples shared a comparison, and retention testing showed a preference for the comparison associated with these two samples, the expected result if the animals used a single-code / default strategy. However, this result is also consistent with the establishment of a code for each sample and a preference for the shortest sample on retention testing, a common result in retention testing with duration samples,

known as choose-short effect (see Grant & Spetch, 1993; Kraemer, Mazmanian & Roberts, 1985; Spetch & Wilkie, 1982, 1983).

To disentangle the choose-short effect from the single-code / default coding, in Study I, we trained pigeons in a task with three nonzero sample durations (2, 6 and 18 s), with the 6- and 18-s samples sharing the correct comparison. In retention testing, the choose-short effect would predict a preference for the comparison associated with the shortest sample (2 s), and single-code / default would predict a preference for the comparison mapped many-to-one, that is, the comparison associated with two samples (6 s and 18 s).

The elements of the environment and the codes derived from those elements are intimately related; they can be thought of as analogous to a recipe and ingredients: The elements of the environment are the "ingredients" the animal can use to establish a set of rules – the "recipe" – that define the relation between those elements and determine which response is controlled by which stimulus, to produce the final result, behavior adapted to the environment. Therefore, to establish adequate response rules, an animal must first identify which stimuli are correlated with important events. For instance, the matching-to-sample tasks discussed previously were designed assuming that the pigeons attend to the sample durations to learn the task. However, in many cases, there are other stimuli that may also be used to perform the task well. Case in point, in Study I the intertrial interval (ITI) was illuminated with a houselight. The end of the ITI and start of the sample were contiguous, so the animal could have learned to use, instead of the onset of the sample keylight, the offset of the houselight as the time marker to initiate timing. Moreover, the offset of the sample keylight was contiguous with the onset of the comparisons, so any of those two events could be used to stop timing. Study II focused on the possibility that the animals could be using the ITI illumination instead of (or in addition to) the samples to learn the task. More specifically, retention and no-sample tests were run to identify which stimuli were used as "start" and "stop" time markers.

The results of Study II were consistent with the pigeons making use of both sample and ITI illumination to learn the task. As a follow-up, Study III aimed to clarify how the two stimuli were being used. In two tests, each of the stimuli was removed to evaluate their relative influences. We also aimed to quantify the relative influence of each stimulus, for each individual bird.

To supplement the previous studies, in Study IV, we used a task with a different sample-comparison mapping: the intermediate sample duration was now the one

mapped one-to-one. Retention and no-sample tests similar to the previous studies were run, to assess whether the different sample-comparison mapping would induce different coding.

Finally, in Study V, to better understand the boundary conditions of the single-code / default strategy, we trained pigeons in a non-temporal discrimination task that shared an equivalent mapping to the previous four temporal-discrimination studies: one color was associated with one comparison, and two other colors were associated with the other comparison. Evidence of the adoption of a single-code / default was found in a non-temporal discrimination that used five samples (Clement & Zentall, 2000), and we tested whether the 3-sample task would yield similar results.

- 6 -

CODING IN PIGEONS: MULTIPLE-CODING VERSUS SINGLE-CODE / DEFAULT STRATEGIES[1]

To investigate the coding strategies that pigeons may use in a temporal discrimination tasks, pigeons were trained on a matching-to-sample procedure with three sample durations (2s, 6s and 18s) and two comparisons (red and green hues). One comparison was correct following 2-s samples and the other was correct following both 6-s and 18-s samples. Tests were then run to contrast the predictions of two hypotheses concerning the pigeons' coding strategies, the multiple-coding and the single-code / default. According to the multiple-coding hypothesis, three response rules are acquired, one for each sample. According to the single-code / default hypothesis, only two response rules are acquired, one for the 2-s sample and a "default" rule for any other duration. In retention interval tests, pigeons preferred the "default" key, a result predicted by the single-code / default hypothesis. In no-sample tests, pigeons preferred the key associated with the 2-s sample, a result predicted by multiple-coding. Finally, in generalization tests, when the sample duration equaled 3.5s, the geometric mean of 2s and 6s, pigeons preferred the key associated with the 6-s and 18-s samples, a result predicted by the single-code / default hypothesis. The pattern of results suggests the need for models that take into account multiple sources of stimulus control.

---

[1] This chapter reproduces the publication: Pinto, C., & Machado, A. (2015). Coding in pigeons: Multiple-coding versus single-code/default strategies. *Journal of the Experimental Analysis of Behavior*, *103*, 472–483.

The ability to learn to behave according to a rule is often referred to as coding (e.g., prospective or retrospective; analogical or non-analogical; for a review, see Grant, Spetch, & Kelly, 1997), and this ability is believed to have been achieved through the evolution of general processes of learning focused on the antecedents and consequents of action (Skinner, 1984). One important goal of research in the field of learning is to understand why and how behavior accords to different rules in different circumstances.

To study coding, researchers have frequently used a delayed matching-to-sample task. In its simplest version, one of two stimuli (samples S1 and S2) is presented and then removed. Afterwards, two other stimuli (the comparisons, C1 and C2) are presented simultaneously. If the animal chooses C1 following S1, or C2 following S2, it receives a reward. Hence, learning the task may be conceived of as learning two conditional discriminations, "If S1, choose C1" and "If S2, choose C2". Of particular interest to test theories and models of coding is the effect on choice accuracy of introducing retention intervals between the samples and the comparisons.

When the samples are stimuli varying in duration (e.g. a light lasting 2 s, S1, or 10 s, S2, the retention functions following S1 and S2 typically diverge. That is, as the retention interval increases, choice following the short sample remains accurate (i.e., the animal continues to prefer C1 over C2), but choice following the long sample becomes increasingly inaccurate (i.e., choice of C2 decreases or, equivalently, choice of C1 increases with the retention interval). This result is known as the choose-short effect (Spetch & Wilkie, 1982; see also Spetch, 1987; Spetch & Wilkie, 1983).

One account of the choose-short effect is the coding model put forth by Kraemer, Mazmanian and Roberts (1985). The model makes four assumptions. First, the animal learns a specific response code for each sample (e.g., "If the light lasts 2 s choose the red key"). Second, during the retention interval, the code is increasingly likely to be forgotten, perhaps because of stimulus interference during the interval and, in some cases, the similarity between the retention interval and the intertrial interval (see Dorrance, Kaiser, & Zentall, 2000; Kelly & Spetch, 2000; Sherburne, Zentall, & Kaiser, 1998). Third, the absence of a code is functionally equivalent to a no-sample or 0-s sample. And fourth, in the absence of a code at the moment of choice, the animal chooses the comparison associated with the sample closest to 0 s.

The model predicts that, as the retention interval increases, the code is increasingly likely to be lost, and therefore the animal is increasingly likely to prefer the comparison associated with the shortest sample (for alternative accounts of the choose-

short effect, see, e.g., Gaitan & Wixted, 2000; Grant, 2009a; Sherburne et al., 1998; Spetch & Wilkie, 1983). Because Kraemer et al.'s (1985) model assumes one code for each sample, it will be referred to as the multiple-coding hypothesis.

In more complex tasks, the multiple-coding strategy may not be the most economical or easiest to learn. Suppose that three samples, S1, S2, and S3, are mapped onto two comparisons, C1 and C2, such that C1 is correct following S1, and C2 is correct following both S2 and S3. In this many-to-one task, instead of learning three codes, the animal could learn a single code for S1 ("If S1, choose C1") and a default rule to be applied following any other sample or even no sample (i.e., "If not S1, choose C2). This seemingly more economical strategy is known as the single-code / default.

To test the single-code / default hypothesis, Singer, Klein, and Zentall (2006, Experiment 2) rewarded pigeons for choosing comparison C1 following 8-s samples, and comparison C2 following 2-s and 32-s samples. After the pigeons learned the task, retention intervals ranging from 1 to 30 s separated the samples from the comparisons (see also Clement & Zentall, 2000, Gaitan & Wixted, 2000). To predict the retention functions, the authors made the following three assumptions. First, in this many-to-one mapping task pigeons learn a single code for 8-s samples and a default rule ("If 8 s, choose C1; otherwise, choose C2"). Second, during the retention interval, the sample code is increasingly likely to be forgotten. And third, in the absence of a sample code at the moment of choice, the pigeon behaves according to the default rule and chooses C2, the comparison associated with the 2-s and 32-s samples. Therefore, the single-code / default hypothesis predicted that, as the retention interval increased, choice following the 2-s and 32-s samples should remain accurate, but choice following the 8-s samples should become increasingly inaccurate. The results confirmed these predictions.

However, Singer et al. (2006)'s findings are not conclusive regarding the animal's coding strategy because they can also be explained by the multiple-coding hypothesis. According to the latter, during training the pigeons learned a specific code for each of the three samples, "If 2 s, choose C2", "If 8 s, choose C1", and "If 32 s, choose C2". During the retention intervals, the codes were increasingly likely to be forgotten and, on those occasions, pigeons chose the comparison associated with the shortest sample, C2. Because C2 was correct following both 2-s and 32-s samples, the accuracy for those samples remained high and the accuracy for the 8-s sample decreased. Therefore, both the single-code / default and the multiple-coding hypotheses account for Singer et al.'s (2006) results.

The present experiment was designed to contrast the two coding hypotheses. To that end, pigeons were exposed to a many-to-one matching task with three samples and two comparisons. Specifically, pigeons learned to choose C1 following 2-s samples, and to choose C2 following 6-s and 18-s samples. With this mapping, the two hypotheses predict different retention function patterns.

According to the multiple-coding hypothesis, forgetting during the retention interval should lead the animal to act as if in a no-sample trial and consequently to choose C1, the comparison associated with the sample closest to 0 s. In terms of retention functions, correct responses following 2-s samples should remain high, whereas correct responses following 6-s and 18-s samples should decrease. The left panel of Figure 1 shows these predictions.

In contrast, if a single-code / default strategy is used, the pigeons should learn a single code for the 2-s samples and apply a default rule in the absence of the 2-s-sample code (i.e., "If 2 s, choose C1, otherwise choose C2"). Hence, as the retention interval increases, the pigeons are more likely to forget the sample code and consequently more likely to choose C2. In terms of retention functions, correct responses following 2-s samples should decrease, whereas correct responses following 6-s and 18-s samples should remain high. The right panel of Figure 1 shows these predictions. Note that, according to both hypotheses, the functions following the 6-s and 18-s samples should not differ.
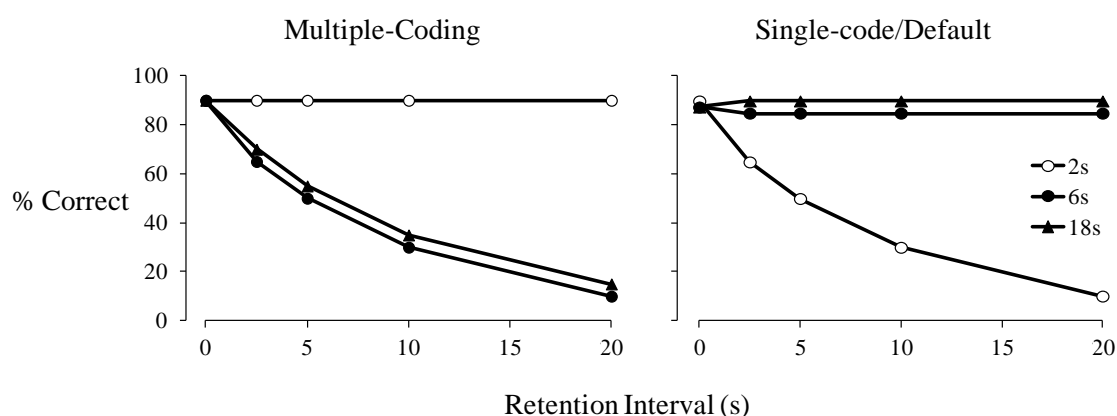
Figure 1. Retention functions predicted by the multiple-coding hypothesis (left panel) and the single-code / default hypothesis (right panel) in a delayed matching-to-sample task with 2-s samples associated with one comparison and 6-s and 18-s samples associated with other comparison.

Another goal of the present study was to test an assumption common to both accounts, namely, that if during the retention interval the memory for the sample (or its code) becomes unavailable, the animal behaves as if in a no-sample trial. To test this assumption, we included test trials without a sample – the comparisons were presented either immediately after the intertrial interval or after a "retention" interval (with no sample, a retention interval is defined procedurally as an interval similar to that included when the trial includes a sample). The two coding hypothesis predict different choice patterns on no-sample trials. Whereas the multiple-coding hypothesis predicts a preference for C1, the comparison associated with the shortest sample, the single-code / hypothesis predicts a preference for C2, the comparison associated with the default rule. Moreover, the preference for C1 (multiple coding) or for C2 (single-code / default) should not vary with the retention interval.

Another reason to manipulate the retention interval on no-sample trials is that the resulting retention function may reveal sources of control over choice other than the sample duration. To illustrate, if choice on no-sample trials without a retention interval differs from the two model predictions, one would conclude that the offset of the intertrial interval also affects choice. Similarly, if choice on no-sample trials changes with the retention interval, then the stimulus conditions introduced by the retention interval (e.g., Dorrance et al., 2000; Grant, 2006b), including the passage of time during the interval, also affect choice. More generally, knowledge of the retention function obtained on no-sample trials may help us to isolate the specific effects of the samples on choice.

The final goal of the present experiment was to obtain in this many-to-one task a psychometric function relating choice proportion to sample duration. We were particularly interested in the bisection point and the overall shape of the psychometric function for their potential implications for the two coding hypotheses.

Consider the bisection point. Previous research has shown that, in one-to-one matching-to-sample tasks with two sample durations and two comparisons, the bisection point occurs at the geometric mean of the two sample durations (e.g., Catania, 1970; Church & Deluty, 1977; Stubbs, 1968, 1976). Although the present task is not one-to-one, it is conceivable that bisection will be determined solely by the 2-s samples (associated with C1) and the 6-s samples (associated with C2); the 18-s samples (also associated with C2) would not affect the bisection point. In this case, the pigeons should be indifferent between C1 and C2 at 3.5 s. But other factors may bias choice, and one of

them is the default rule: If 3.5-s samples are sufficiently distinct from 2-s samples to activate the single code, the default rule will be triggered and choice following 3.5-s samples will be biased toward C2. If pigeons prefer C2 at the geometric mean of 2s and 6s, then they will be indifferent at a duration shorter than the geometric mean.

Consider now the shape of the psychometric function. According to the multiple-code hypothesis, when a sample code is lost the animal chooses the comparison associated with the sample closest to 0 s, in our case, the 2-s sample. It follows that preference for C1 should be strong and roughly constant for samples in the range of 0 to 2 s. Moreover, if we make the reasonable assumption that each code generalizes to sample durations around the trained sample duration, then the multiple code hypothesis predicts a typical psychometric function, with a bisection point at the geometric mean of 3.5 s. The results of the temporal generalization tests may help us decide among these different possibilities.

The two comparison keys can be labeled in two different ways: as a function of their associated sample duration (the correct key following 2-s samples [C1] is the "short" key and the correct key following 6-s and 18-s samples [C2] is the "long" key) or as a function of their meaning according to the single-code / default hypothesis (the correct key following 2-s samples [C1] is the "single-code" key and the correct key following 6-s and 18-s samples [C2] is the "default" key). These two nomenclatures will be used hereafter.

Method

Subjects

Six pigeons (*Columba livia*) maintained at approximately 80% of their free-feeding body weight served as subjects. Water and grit were freely available in their home cages. The pigeon room was maintained in a 13:11 hour light/dark cycle, with the lights on at 08:00, and its temperature was kept between 20º-22º C. The experiment was conducted once a day, 7 days a week, at approximately the same time of day for each pigeon.

Three of the pigeons (P463, P501, and P536) had previously participated on a timing experiment, although with comparison stimuli different from the ones used in this experiment, and the three other pigeons (P785, P917, P973) were experimentally naive.

Apparatus

Six operant chambers were used: Five LVE (Lehigh Valley Electronics) chambers, and a homemade chamber. The LVE chambers measured 34 x 35 x 31 cm (h x l x w). Three circular response keys, 2.5 cm in diameter, were arranged horizontally on the response panel. The bottom edge of each key was 22.5 cm above the wire mesh floor, and the keys were 9 cm apart, center to center. Each key was equipped with a 12-stimulus IEE (Industrial Electronics Engineers) in-line projector. The food hopper was accessible through a 6-cm wide x 5-cm high opening that was centered horizontally on the response panel, 8.5 cm above the floor. When the hopper was raised, a 28-V, 0.04-A light illuminated its opening and grain became accessible to the pigeon. On the wall opposite the response panel, 30 cm above the floor, a 28-V, 0.1-A houselight provided general illumination. The operant chamber was enclosed in an outer box equipped with an exhaust fan. The fan circulated air through the chamber and masked outside noises.

The homemade chamber measured 31 x 33 x 33 cm (h x l x w). Three circular response keys, 2.5 cm in diameter, were arranged horizontally on the response panel. The bottom edge of each key was 21 cm above the wire mesh floor, and the keys were 9 cm apart, center to center. Each key was equipped with a 12-stimulus IEE in-line projector. A LVE food hopper was accessible through a 6-cm wide x 4.5-cm high opening that was centered horizontally on the response panel, 6.5 cm above the floor. When the hopper was raised, a 28-V, 0.04-A light illuminated its opening and grain became accessible to the pigeon. On the wall opposite the response panel, 27.5 cm above the floor, a 28-V, 0.1-A houselight provided general illumination. The operant chamber was enclosed by a PVC sound-attenuating cubicle (Med Associates, ENV-018V) equipped with an exhaust fan.

In this experiment, the side keys were illuminated with red or green hues and the central key was illuminated with a white hue. Personal computers using the ABET II (Lafayette Instrument Company) software controlled the experimental events and recorded the data.

Procedure

Training. The birds were trained in a symbolic matching-to-sample task. Following the presentation of a white hue on the center key for 2, 6 or 18 s (sample stimulus), each of the side keys was illuminated with either a red or a green hue (comparison stimuli). One of the comparisons was correct following the 2-s sample, and

the other comparison was correct following the 6-s and 18-s samples (the correct comparison for each sample was counterbalanced across pigeons). After a response, the comparison keys were turned off and, if the response was correct, reinforcement was delivered and the 30-s intertrial interval (ITI) started. If the response was incorrect, the ITI started immediately. To avoid confusion between the ITI and the dark retention interval used in subsequent tests, the houselight was illuminated during the ITI but was turned off at sample onset. A correction procedure was in effect: Following an incorrect response, the trial was repeated; after three consecutive incorrect responses, only the correct comparison key was presented. The birds began each session at approximately 80% of their free-feeding weight and the reinforcement duration was adjusted for each bird to minimize feeding outside the experimental session. It varied from 1.5 to 4.5 s across animals.

Each session comprised sixty-four trials (excluding correction trials), thirty-two 2-s sample trials, sixteen 6-s sample trials, and sixteen 18-s sample trials. Across trials, the location of the comparison stimuli varied pseudo-randomly with the constraint that each comparison stimulus was presented the same number of times on each of the side keys. Training continued until the pigeon met a criterion of at least 80% correct responses to each sample in a session (excluding correction trials), for five consecutive sessions, or until 40 sessions were completed.

Retention Test. After the training phase, a retention interval was introduced between the sample and comparison stimuli. The retention interval – spent in darkness – could be 2.5, 5, 10 or 20-s long. Each session comprised 80 trials, 48 regular training trials (24 x 2 s, 12 x 6 s, 12 x 18 s) and 32 retention-interval test trials (16 x 2 s, 8 x 6 s, 8 x 18 s). Irrespective of trial type, correct responses were reinforced, but the correction procedure was in effect only on regular trials (with no retention interval). To minimize feeding outside the experimental session, the reinforcement durations were recalculated and varied from 1 to 4 s across animals. Testing continued for 30 sessions.

Retraining I. The birds returned to a training phase until they made at least 80% correct responses to each sample (excluding correction trials) for 5 consecutive sessions or until 15 sessions were completed.

No-sample Test. In this test, no-sample trials were interspersed among the regular training trials. In a no-sample trial, the comparison keys were presented immediately after the ITI. Each session comprised 72 trials, 64 regular training trials (32 x 2 s, 16 x 6 s, 16 x 18 s) and 8 no-sample test trials. Responses on no-sample trials

were never reinforced. To minimize feeding outside the experimental session, the reinforcement durations were recalculated and varied from 1.5 to 5 s across animals. Testing continued for 10 sessions.

Retraining II. This phase was exactly the same as Retraining I.

Generalization Test. Two ranges of untrained sample durations were presented. The first range included samples of 1, 3.5, 10.4 and 36 s. Two of these values (1 s and 36 s) were outside the training range. The other two (3.5 s and 10.4 s) were inside the training range and were equally discriminable from the adjacent training durations (3.5 s is the geometric mean of 2 s and 6 s, and 10.4 s is the geometric mean of 6 s and 18 s). The second range included samples of 3, 4, 5 and 10.4 s. The range spanned the interval with higher variability during the first test; the duration of 10.4 s was common to both test ranges and allowed a direct comparison between them.

Each session comprised 96 trials, 56 regular training trials (28 x 2 s, 14 x 6 s, 14 x 18 s) and 40 generalization trials (10 trials for each of test sample). The session was divided into six blocks of 16 trials each. The first block included only regular trials (8 x 2 s, 4 x 6 s, 4 x 18 s). The next five blocks included eight regular trials (4 x 2 s, 2 x 6s, 2 x 18 s) and eight generalization trials (two trials per test sample). Responses following test samples were not reinforced. To minimize feeding outside the experimental session, the reinforcement durations on regular trials were recalculated and varied from 2 to 6 s across animals. Testing continued for 10 sessions, 5 for each range.

Retraining III. This phase was exactly the same as Retraining I except that the maximum number of session was reduced to 10.

No-sample Retention Test. This test was similar to the No-sample Test, with the exception that, on some of the no-sample trials, a retention interval of 2.5, 5, 10 or 20 s was introduced. Therefore, at the end of the ITI the houselight was turned off for the duration of the retention interval and then the comparison keys were illuminated. Each session comprised 78 trials, 48 regular training trials (24 x 2 s, 12 x 6 s, 12 x 18 s) and 30 no-sample trials. Of the no-sample trials, 6 had no retention interval and 24 had a retention interval (6 trials for each of the 4 retention interval durations). No-sample trials were never reinforced. To minimize feeding outside the experimental session, the reinforcement durations on regular trials were recalculated and varied from 2 to 6 s across animals. Testing lasted 10 sessions.

Results

Training & Retraining. Five of the six pigeons met the learning criterion in 13 to 29 sessions (average of 20 sessions). The exception, P917, although failing to reach criterion in 40 sessions, acquired the discrimination (its matching accuracy during the last 5 training sessions equaled 78% for 2-s samples, 84% for 6-s samples and 99% for 18-s samples).

Due to an equipment malfunction, when switching to the Retention Test, three of the birds were exposed to sessions where the ITI was not illuminated. P463 ran one session in such conditions, P501 ran three sessions and P536 ran seven sessions. These birds returned to the training phase until the criterion was reached again or a total of 40 training sessions was completed. P463 ran five additional training sessions, P536 ran six additional sessions, and P501 failed to reach the criterion, having run 20 additional sessions (until the maximum of 40 sessions was reached). P501's failure to reach the criterion was due to the 6-s samples: Its matching accuracy for the last 5 training sessions equaled 88% for 2-s samples, 76% for 6-s samples and 99% for 18-s samples. Although failing to reach the criterion, P501 continued to next phase of the experiment.

The birds needed from 7 to 15 sessions ($\bar{x} = 12$) to complete Retraining I, from 5 to 15 sessions ($\bar{x} = 8$) to complete Retraining II and from 5 to 10 sessions ($\bar{x} = 6$) to complete Retraining III.

Retention Test. In the Retention Test there were two types of trials: training trials with no retention interval, and retention-test trials. On both trials, correct responses were reinforced. Therefore, reinforcement on test trials could have changed performance during testing. To assess whether performance changed during testing, we analyzed separately the data from the first test session and from all 30 test sessions.

Figure 2 shows the results of the first test session. Matching accuracy on 2-s trials decreased abruptly with the shortest retention interval and then stabilized, whereas matching accuracy on 6-s and 18-s trials did not decrease as much and as abruptly. A two-way repeated-measures ANOVA with sample duration (three levels) and retention interval (five levels) as factors revealed a significant main effect of sample duration, $F(2, 10) = 10.91$, $p = .003$, $\eta_G^2 = .45$, and of retention interval, $F(4, 20) = 16.60$, $p <.001$, $\eta_G^2 = .27$. The interaction also was significant, $F(8, 40) = 2.54$, $p = .024$, $\eta_G^2 = .18$, confirming that the retention interval did not affect matching accuracy equally

following the three samples. This pattern of results is consistent with the single-code / default hypothesis.

Figure 2 also shows that percent correct following 6-s and 18-s samples decreased with retention interval (a repeated-measures ANOVA showed a significant effect of retention interval, $F(4,20) = 4.53$, $p = .009$, $\eta_G^2 = .17$). Correct responses following 6-s samples seemed to be below correct responses following 18-s samples, but that difference was not significant: $F(1,5) = 1.29$, $p = .301$, $\eta_G^2 = .05$). Moreover, the decrease in percent correct following the 2-s samples occurred abruptly from no retention interval to the 2.5-s retention interval, but it did not change with longer intervals (a repeated-measures ANOVA showed no significant effect of non-zero retention interval on 2-s samples, $F(3, 15) = .909$, $p = .460$, $\eta_G^2 = .06$). Finally, on retention trials, percent correct following 2-s samples was significantly below indifference (95% Confidence Interval = 20% - 38%).
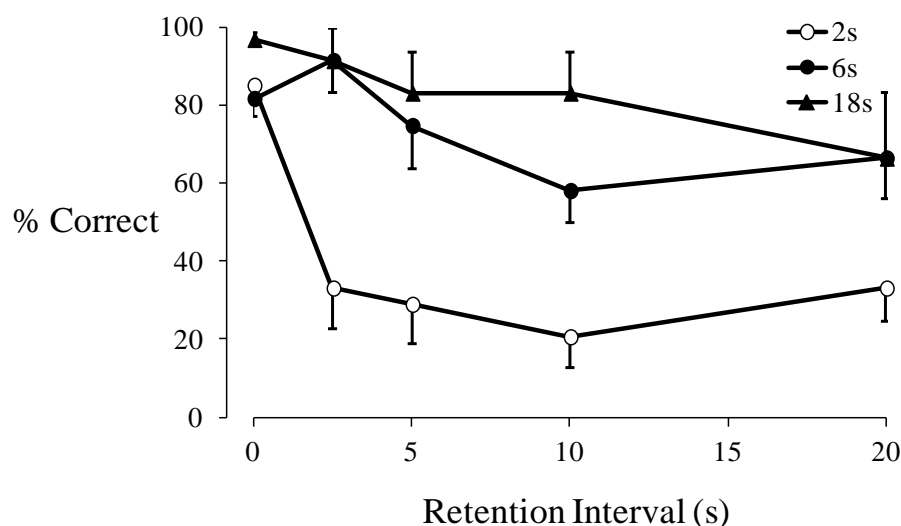
Figure 2. Mean (with SEM) percent correct to each of the three sample durations as a function of retention interval duration. The data come from the first test session of the Retention Test.

Data from all test sessions were divided into six 5-session blocks. Figure 3 shows data from the first and the last blocks. The results from the first block were similar to the first session (compare the left panel of Figure 3 with Figure 2): A two-way repeated-measures ANOVA with sample duration (three levels) and retention interval (five levels) as factors revealed a significant interaction, $F(8, 40) = 2.65$, $p = .020$, $\eta_G^2 = .22$, confirming that, during retention testing, matching accuracy to the three samples evolved differently. Similarly to the first session, percent correct following 6-s and 18-s

samples decreased with retention interval (a repeated-measures ANOVA showed a significant effect of retention interval, $F(4,20) = 9.98$, $p < .001$, $\eta_G^2 = .41$). Additionally, matching accuracy to 6-s samples was lower than to 18-s samples (significant main effect of sample: $F(1,5) = 13.29$, $p = .015$, $\eta_G^2 = .20$). However, as the test progressed, the response pattern changed and the three curves approached each other (right panel, Figure 3). By the end of testing, the curves for all three sample durations approached indifference as the retention interval increased.
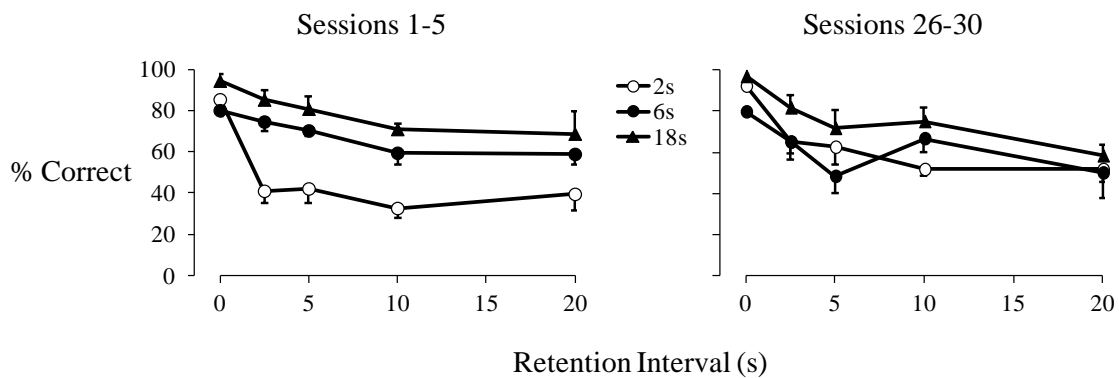
Figure 3. Mean (with SEM) percent correct to each of the three sample durations as a function of retention interval in the first 5 sessions (left panel) and last 5 sessions (right panel) of the Retention Test.

To understand the evolution of responding during testing, Figure 4 shows percent correct on test trials (trials with a retention interval) across the six blocks, with sample duration as the parameter. As testing progressed, there was an increase in correct responses following the 2-s samples, and a slight decrease in correct responses following the 6-s. The net result was the maintenance of average percent correct across blocks (see dotted line). A two-way repeated-measures ANOVA with block (six levels) and sample duration (three levels) as factors confirmed this interpretation: No main effect of block was found, $F(5, 25) = .794$, $p = .564$, $\eta_G^2 = .02$, but there was a significant main effect of sample duration, $F(2, 10) = 9.48$, $p = .005$, $\eta_G^2 = .43$, and of the interaction, $F(10, 50) = 2.95$, $p = .005$, $\eta_G^2 = .16$. Accuracy on test trials did not evolve similarly following each sample.
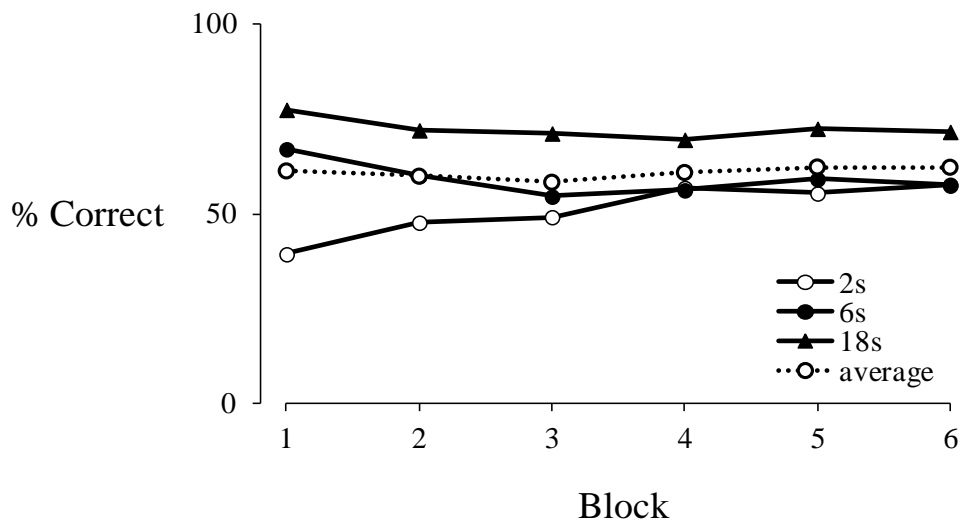
Figure 4. Mean percent correct on test trials of the Retention Test to each of the three sample durations as a function of testing sessions (each block is composed of 5 sessions). The dotted line is the average of the three samples.

No-sample Test. Figure 5 shows the results for the No-sample Tests, with and without retention intervals. In tests with no (or 0-s) retention intervals, most choices were to the "short" key (%Long = 28% and 32% in the two phases, a non-significant difference, $F(1, 5) = .47$, $p = .525$, $\eta_G^2 = .05$). Moreover, a t-test showed that %Long differed significantly from 50%, $t(5) = 5.42$, $p = .003$, $d_z = 2.21$, 95% CI for d [0.64, 3.74]. Finally, at the individual level, a two-tailed normal approximation to the binomial showed that preference for "Short" was below chance in five of the six birds ($\alpha = .05$). These results are consistent with the multiple-coding hypothesis.

The data from the No-sample Test with retention intervals show that, as the retention interval increased, the percentage of "long" responses converged to chance. At the longest retention interval (20s), overall preference for the "long" key had clearly increased: A two-tailed normal approximation to the binomial ($\alpha = .05$) showed that, while two birds continued to prefer the "short" key, two preferred the "long" key, and two did not differ significantly from chance. Neither the multiple-coding nor the single-code / default hypotheses predicted this result.
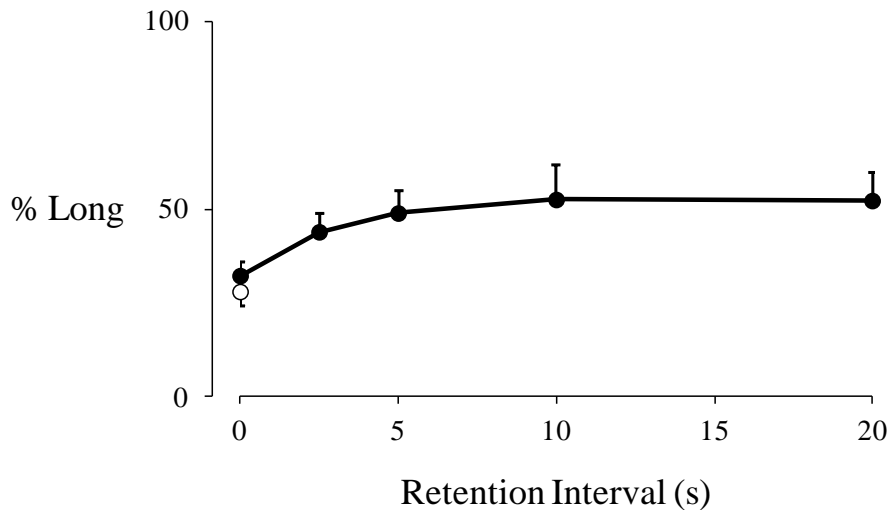
Figure 5. Mean (with SEM) percent of choices to the "long" key (key associated with 6-s and 18-s samples) as a function of retention interval on the No-sample Test. The white dot is the result of the first No-sample Test.

Generalization Test. Two ranges of samples were presented. To test whether performance differed between the two ranges, performance on samples common to both ranges (2, 6, 10.4 and 18 s) was compared via a two-way repeated-measures ANOVA with test range (two levels) and sample duration (four levels) as factors. There were no significant effects of test range ($F(1, 5) = .76$, $p = .424$, $\eta_G^2 = .01$) or its interaction with sample duration ($F(3, 15) = 1.37$, $p = .290$, $\eta_G^2 = .07$). Therefore, we combined the results from the two ranges and averaged the data from the common samples.

Figure 6 shows the mean percent of "long" or "default" choices as a function of sample duration. The general pattern of the psychometric function was the one expected on the basis of temporal generalization: When a sample was presented, preference for the "long" key increased with sample duration according an ogive function. As for choice percentage following the 3.5-s samples (the geometric mean of 2 s and 6 s), the multiple-coding hypothesis predicted indifference, whereas the single-code / default hypothesis predicted a preference for the "default" key. The result, 73% choices for the "long" or "default" key, was consistent with the latter (a t-test showed that choices differed significantly from 50%, $t(5) = 3.96$, $p = .011$, $d_z = 1.62$, 95% CI [0.33, 2.85]). At the individual level, a two-tailed normal approximation to the binomial showed that the preference for the "long" or "default" key was significantly above chance in four of the six birds.

Finally, it is also worth noting that preference following 0-s samples was significantly different from preference following the durations nearest to it: 1-s samples ($F(1, 5) = 43.60$, $p = .001$, $\eta_G^2 = .77$) and 2-s samples ($F(1, 5) = 33.57$, $p = .002$, $\eta_G^2 = .73$).
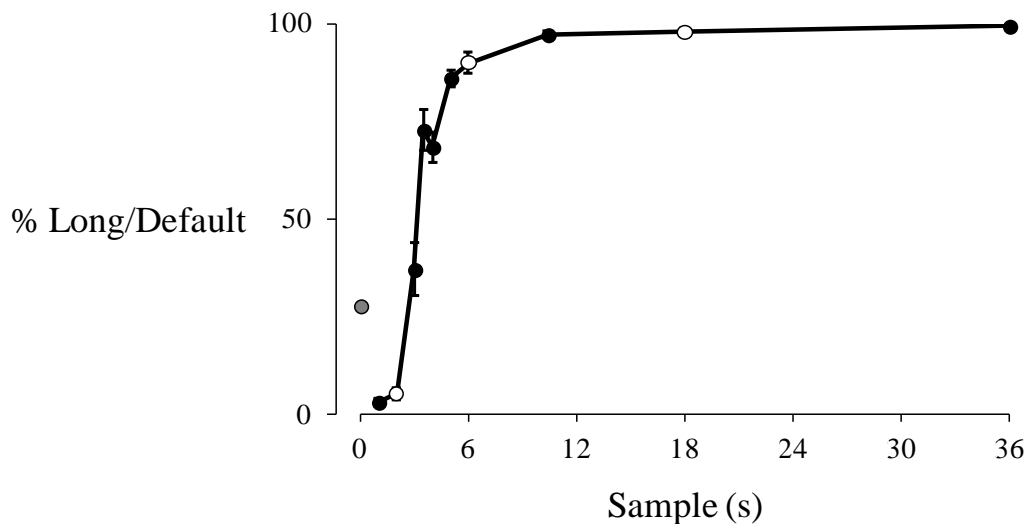
Figure 6. Mean (with SEM) percent of choices to the "long" or "default" key (key associated with 6-s and 18-s samples) as a function of sample duration on the Generalization Test. The white dots identify the previously-trained durations (2 s, 6 s and 18 s) and the gray dot is the result of the first No-sample Test.

Discussion

This study attempted to uncover the coding strategies adopted by pigeons in a matching-to-sample task where one comparison was correct after one sample and another comparison was correct following two samples. The pigeons could adopt one of two strategies, use three codes, one for each sample (multiple-coding hypothesis), or use two codes, one specific to the 2-s sample, and a default code triggered by any other sample (single-code / default hypothesis).

We examined the two hypothesis in the light of three types of tests, a retention-interval test, a no-sample test (with and without a retention interval), and a generalization test. Trials with a retention interval tested whether, when sample information is lost, pigeons prefer the "short" key, as the multiple-coding hypothesis proposes, or the "default" key, as the single-code / default hypothesis proposes. Trials without a sample tested whether such trials are functionally equivalent to losing sample

information during a retention interval, an assumption shared by both hypotheses. Generalization trials tested the predictions of each hypothesis concerning the location of the bisection point, at the geometric mean according to the multiple-code hypothesis, or below the geometric mean according to the single-code / default hypothesis.

The general pattern of results found in the Retention Test supported the single-code / default hypothesis: When retention intervals were introduced, the birds showed a preference for the "default" key. However, some results are hard to reconcile with the hypothesis. First, percent correct following the 6-s and 18-s samples decreased with retention interval (see Figure 2 and left panel of Figure 3). Since on these trials a 2-s sample was not presented, according to the hypothesis, the pigeons should have chosen the "default" key, both on trials with and on trials without a retention interval. Hence, percent correct following the 6-s and 18-s samples should not decrease, as it did. Second, percent correct on the 6-s samples was generally below percent correct on the 18-s samples. Since 6-s and 18-s samples share the same "default" response, the single-code / default hypothesis does not predict a difference between these two functions. Third, also not predicted by the single-code / default hypothesis was the abrupt decrease in accuracy on the 2-s sample trials following the shortest retention interval, accuracy that then remained relatively stable with longer intervals. A progressive decrease would be expected given that the retention interval should increase the probability of losing the sample code, and therefore increase the probability of choosing the "default" key. These discrepancies suggest that the single-code / default hypothesis needs to be elaborated with additional principles to account for the full range of effects of the retention interval.

Pinto and Machado (2011) suggested that multiple effects may be present in delayed matching-to-sample tasks. One of them is the disruption of timing and consequent random responding produced by the retention interval. Specifically, the longer the retention interval, the more likely the disruption and the closer to indifference choice should be. This effect could explain the decrease in matching accuracy following the 6-s and 18-s samples (Figure 2 and Figure 3). Further evidence consistent with the effect can be seen in the No-sample Tests (Figure 5): As the retention interval increased, choice also approached chance.

To further clarify the retention interval effect, Figure 7 re-plots the percentage of "long" choices as a function of sample duration (including no, or 0-s, samples), with the retention interval as a parameter. The data for the 2-s, 6-s, and 18-s samples come from

the Retention Test trials, and the data for the 0-s samples come from the No-sample Test trials. The curve for trials without a retention interval (filled triangles) shows a preference for the "short" key following 2-s samples, and a preference for the "long" key following 6-s and 18-s samples. This result is expected because these were the choices the birds learned during training. However, when retention intervals were introduced (circles), the curve flattened and approached indifference. In fact, the longer the retention interval, the closer the curve came to indifference (contrast the empty and filled circles).

Another effect that could account for some results at odds with the single-code / default hypothesis is stimulus generalization. This effect may explain why performance on 18-s samples was generally better than on 6-s samples, even though they shared the same comparison key (see Figure 3). Because the closer two stimuli are, the harder it is to discriminate between them, the discrimination between the 2-s and 6-s samples may have been more difficult than the discrimination between the 2-s and 18-s samples. Hence, percent correct following the 6-s samples was not as good as that following the 18-s samples.
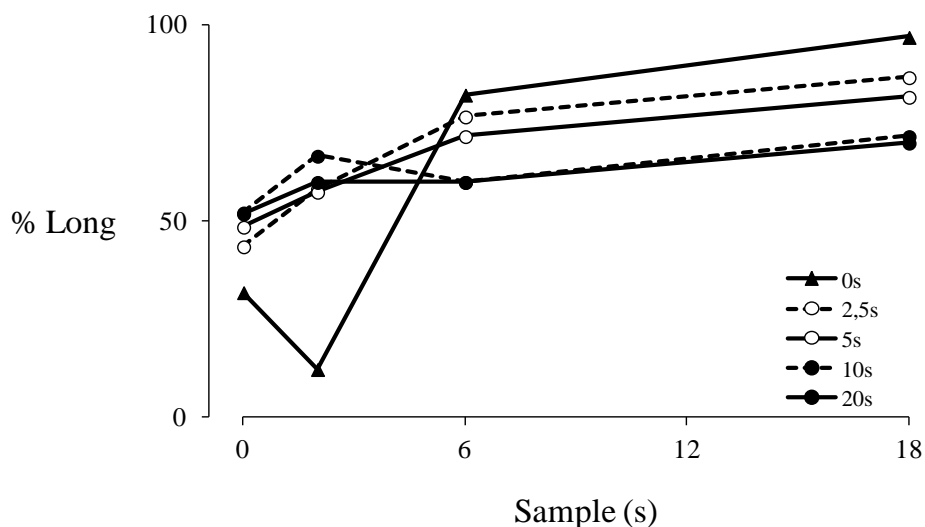
Figure 7. Mean percent of choices to the "long" key (key associated with 6-s and 18-s samples) as a function of the sample duration presented. Each line refers to a retention interval. Data points for 2-s, 6-s, and 18-s samples come from the Retention Test and data points for 0-s samples come from the No-sample Test.

Data from the Generalization Test trials lends further support to this hypothesis. The difference in correct choices between 6-s and 18-s samples (Figure 6) suggests that

the 6-s samples may have been coded as "short" slightly more often than the 18-s samples. If that was indeed the case, then it would follow that performance on the Retention Test trials would be slightly worse following the 6-s samples than following the 18-s samples. In fact, the average difference in correct choices between the 6-s and 18-s samples on the Generalization Test trials (Figure 6) was of the same order (around 10%) than the average difference in performance between the 6-s and 18-s samples on the Retention Test trials (left panel of Figure 3).

Finally, interpreting how matching accuracy following the 2-s samples changed with retention interval is challenging. According to the single-code / default hypothesis, we should expect a *gradual* decrease. The data showed an abrupt decrease, which did not seem to be affected by the duration of the retention interval (see Figure 2 and left panel of Figure 3). Perhaps generalization decrement due to stimulus changes brought about by the retention interval (dark period) rather than the duration of a retention interval causes the forgetting.

Yet another possibility to explain the pattern of responding following the 2-s samples combines two effects, a preference for the "default" key (as the single-code / default hypothesis predicts) and the aforementioned tendency for performance to approach chance with increasing retention intervals. The former effect would "push" the percent correct function toward 0%, while the latter would "pull" it toward 50%. Percent correct following 2-s samples would decrease until the two forces reached equilibrium and then it would stabilize around a value between 0% and 50%. The same two forces acting on performance following the 6-s and 18-s samples would "push" percent correct to 100% and "pull" it toward 50%. The resulting force would maintain percent correct between these values, decreasing slightly because the pull to 50% would get stronger with the retention interval. In conclusion, the overall data set seems to require the integration of different effects.

The results of the No-sample Test were in the direction predicted by the multiple-coding model, that the pigeons would choose the comparison associated with the sample closest to 0 s. But if the preference for the "short" key following a 0-s sample were due to temporal generalization, then in the Generalization Test (Figure 6), we would expect similar accuracies following the 0-s, 1-s, and 2-s samples. That was not the case. Following 1-s and 2-s samples, only 3% and 5%, respectively, of choices were to the "long" key, but following the 0-s samples, 28% of choices were to the "long" key. Although "short" remained the preferred key following the three samples,

the preference following the 0-s sample was not as strong as expected from the generalization gradient alone. This result suggests a qualitative difference between zero and non-zero samples, perhaps the effect of generalization decrement due to stimulus changes rather than sample duration, an effect pushing performance toward indifference.

It is worth noting that a result based on generalization would not be incompatible with a single-code / default strategy, if we assume that a trial with a 0-s sample could be viewed as functionally similar to a trial with a 2-s sample. In that case, on 0-s trials, the birds would respond according to the "single-code" rule and choose the 2-s key.

Even though the no-sample test by itself may not be conclusive in telling us what coding strategy was in use, its results are informative in regards to an assumption shared by both models: the loss of sample information during a retention interval is functionally equivalent to a 0-s sample trial. If that were the case, we would expect similar preferences following a long retention interval and following a 0-s sample. The results were inconsistent with this prediction. In the Retention Test, the pigeons preferred the 6-s and 18-s key, but in the No-sample Test they preferred the 2-s key. Moreover, preference following the 0-s samples varied in an orderly fashion with the retention interval – as the interval increased, preference approached indifference. We conclude that, contrary to both models, the loss of sample information during a retention interval is not equivalent to a 0-s sample.

A final piece of evidence consistent with the single-code / default hypothesis was the bisection point of the psychometric function obtained during the Generalization Test (Figure 6). The hypothesis predicted the obtained result, a preference for "long" at the geometric mean of 2s and 6s or, equivalently, a bisection point slightly below the geometric mean. Hence, the default rule seems to determine choice following sample durations sufficiently away from 2s.

In addition to the multiple-code and single-code / default, our results suggest a third strategy. The pigeons could have timed the interval from the end of the ITI to the beginning of the choice period, and then compared its duration with a threshold set between 2 s and 6 s; if the interval was below the threshold, they chose the "short" comparison; if above, they chose the "long" comparison. In this strategy, the effective time marker is not the onset of the center key light, the nominal time marker, but a more salient event, the offset of the houselight; and the effective sample is not the interval

during which the center key is illuminated, the nominal sample, but the interval since the houselight was turned off.

This third strategy could explain the results of the Retention Tests. On those trials with both the nominal sample and a retention interval, the effective sample would almost always exceed the threshold, leading the pigeons to prefer the "long" comparison. Therefore, matching accuracy would not change with the retention interval following the 6-s and 18-s nominal samples, but it would decrease abruptly following the 2-s nominal samples – see Figure 2. The strategy could explain also the results of the Stimulus Generalization Tests. The sigmoid curve in Figure 6, with a steep slope at the indifference point, is consistent with a threshold-based account.

However, the strategy is hard to reconcile with the results of the no-sample tests. Pigeons should prefer the "short" comparison more strongly following 0-s nominal samples than 2-s nominal samples because the former should lead more than the second to effective samples below threshold. This result was not observed.

Similarly, if the pigeons followed the third strategy they should have preferred the "long" comparison when the 0-s nominal samples were followed by long retention intervals. This result also did not occur (e.g., in Figure 5, pigeons were indifferent between the comparisons at the 20-s retention interval).

Although this third strategy is plausible (see also Spetch & Rusak, 1989, 1992, for the effects of the ITI on matching-to-sample performance), it does not account for all of our main findings. Future work should explore the possibility that pigeons use multiple time markers (e.g., houselight offset, center keylight onset) and time multiple intervals.

In conclusion, of the two hypotheses put to test, the single-code / default provided more accurate predictions. However, neither hypothesis was consistent with the overall pattern of results. We argued that the pattern may result from a combination of different effects. Identifying the causal processes operating in matching-to-sample tasks and how these processes interact is fundamental to improve our understanding of how animals behave in environments where different coding strategies are possible.

- 28 -

STUDY II

UNRAVELING SOURCES OF STIMULUS CONTROL IN A TEMPORAL DISCRIMINATION TASK[2]

In temporal discriminations tasks, more than one stimulus may function as a time marker. We studied two of them in a matching to sample task, the sample keylight and the houselight that signaled the intertrial interval. One group of pigeons learned a symmetrical matching-to-sample task with two samples (2 s or 18 s of a center keylight) and two comparisons (red and green side keys), whereas another group of pigeons learned an asymmetrical matching-to-sample task with three samples (2 s, 6 s and 18 s) and two comparisons (red and green). In the asymmetrical task, 6-s and 18-s samples shared the same comparison. In a subsequent retention test, both groups showed a preference for the comparison associated with the longer samples, a result consistent with the hypothesis that pigeons based their choices on the duration elapsed since the offset of the houselight (i.e., sample duration + retention interval). Results from two no-sample tests further corroborated the importance of the ITI illumination as a time-marker: When the ITI was illuminated, the proportion of choices correlated positively with the retention interval; when the ITI was darkened, choices fell to random levels. However, the absolute value of choice proportions suggested that the sample stimulus was also a time marker. How multiple stimuli acquire control over behavior and how they combine remains to be worked out.

In temporal discrimination tasks animals learn to behave according to the duration of one of more stimulus intervals. However, it is not always clear which stimuli signal the onset or the offset of the timed interval, that is, which stimuli function as time-markers. To illustrate, consider a matching-to-sample task for pigeons. On each trial, a center keylight is illuminated for 2 s or 10 s (sample) and then the pigeon chooses between a green key and a red key (comparisons), with green correct after the 2-s sample, and red after the 10-s sample. To choose the correct comparison, an animal needs to identify the aforementioned rules governing the task, and to behave according to them. Researchers usually assume that sample onset and offset mark, respectively, the beginning and end of the duration that controls the pigeon's choices. However, other stimuli may also become time markers. For instance, the events that precede the trial (e.g., the intertrial interval, or ITI) can influence the birds' performance on matching tasks. This result may be revealed by inserting a retention interval between sample and comparisons. When the ITI and the retention interval are physically similar, pigeons' preference for the comparison associated with the shortest sample increases with the retention interval, the choose-short effect (e.g., Grant & Spetch, 1993; Spetch & Wilkie, 1982; 1983). But when the ITI and the retention interval are dissimilar, the choose-short effect usually is either weakened or eliminated (Pinto & Machado, 2011; Sherburne, Zentall, & Kaiser, 1998; cf. Kelly & Spetch, 2000). The events that precede the trial influence how the pigeon responds to the sample duration.

In the same vein, changing the ITI duration also can affect choice in temporal discrimination tasks. Spetch and Rusak (1989) found that, when tested with ITIs shorter than in training, pigeons were biased towards the "long" comparison; when tested with ITIs longer than in training, they were biased towards the "short" comparison (see also Spetch & Rusak, 1992). These examples further attest that temporal discrimination may depend on stimuli other than the sample.

Even studies not directly concerned with time-markers have shown the influence of the ITI on temporal discrimination. In Study I, we were interested in pigeons' coding strategies in asymmetrical matching-to-sample tasks. During training, the pigeons learned to map three samples (2 s, 6 s and 18 s) onto two comparisons (red and green hues) such that one comparison – say, Red – was correct

following the 2-s samples and the other comparison, Green, was correct following both the 6-s and 18-s samples. The aim was to test whether the pigeons' performance accorded to a single-code / default rule, a single code for the 2-s sample, and a default rule for all other samples (see also Clement & Zentall, 2000; Singer, Klein, & Zentall, 2006). Since the single-code would be triggered by the shortest sample, mapped to Red, the default response rule would be triggered whenever the sample was not 2-s long, i.e., "not short", and it would involve the choice of Green, the correct choice following the 6-s and 18-s samples. In sum, the single-code / default rule would be "If 2 s, choose Red; otherwise, choose Green". The rule predicts that during a retention interval, as sample information is increasingly likely to be forgotten, the pigeons are more likely to use the default rule and therefore choose the "long" comparison – a choose-long effect. When we introduced retention intervals ranging from 2.5 s to 20 s, the pigeons showed a preference for the "long" comparison after all retention intervals.

Although the result is consistent with a single-code / default strategy, it is also consistent with the hypothesis that, instead of the sample onset, the stimulus signaling the end of the ITI functioned as the effective time marker. In Study I, the 30-s ITI was illuminated by the houselight; the sample was the center keylight; and the retention interval was spent in darkness. During training, the end of the ITI, signaled by the houselight offset, and the start of the sample, signaled by the center keylight onset, occurred simultaneously, and the same was true for the sample offset, signaled by the center keylight offset, and the comparisons onset, signaled by the onset of the red and green keylights. If the pigeons used the houselight offset[3] to initiate timing, perhaps because the houselight is more salient than the center keylight, they could have timed the interval from the houselight offset until the comparisons onset. Because this interval equals the sample duration during training, we cannot tell which time marker is effectively being used. However, the 'ITI offset-comparisons onset' hypothesis predicts that with retention intervals, the effectively timed durations should increase and therefore the pigeons should prefer the "long" comparison, the observed choose-long effect (see bottom panel of Figure 8, Time since ITI > Sample duration).

---

[3] The animals could be timing the interval since the houselight was turned on (start of the 30-s ITI) or turned off (end of the ITI). Seeing that durations since houselight onset would be harder to discriminate (32 s vs. 36 s vs. 48 s) than durations since houselight offset (2 s vs. 6 s vs. 18 s), we assumed that houselight offset was the time marker most likely to be used.
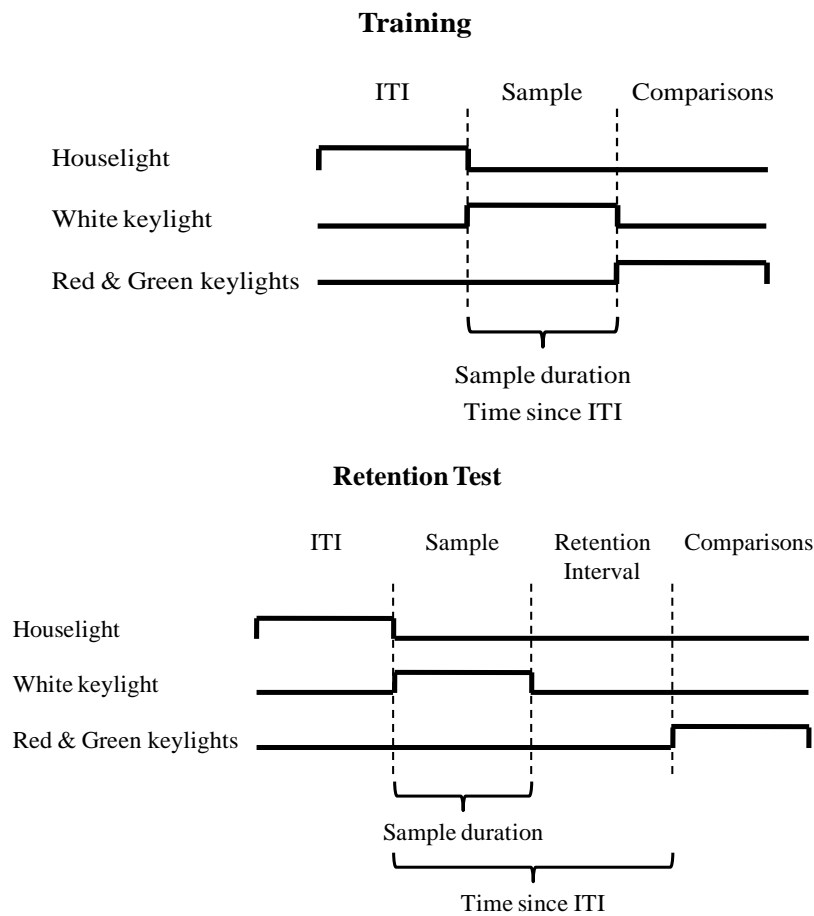
Training

Retention Test

Figure 8. Schematic of the stimuli presented during a trial in Training (top panel) and Retention Tests (bottom panel). Each line is raised whenever its corresponding stimulus was turned on.

This alternative account is also consistent with the results obtained in studies on the timing of empty intervals, that is, intervals signalled by brief start and stop markers (Grant, 2001; Grant & Talarico, 2004; Santi, Hornyak, & Miki, 2003; Santi, Ross, Coppa, & Coyle, 1999). In these studies, the subjects seem to add the retention interval to the sample stimulus, leading to a "long"-key bias.

Identifying the exact time marker is relevant to the study of coding because it clarifies the sources of information an animal uses to establish its response rules. Additionally, knowing which stimulus is being used can change how researchers interpret the results of common tests. For instance, the interpretation of a retention test result changes if an animal is timing the duration of the sample stimulus or the duration of the interval since the end of the ITI. In the former case, the sample and

retention interval durations remain different and separate; in the latter case, the sample and retention interval durations are added.

In the present work, we aimed to clarify which stimuli control responding when multiple stimuli can be used as time markers (as is true for many timing tasks). To that end, one group of pigeons learned a task with two samples (2 s and 18 s) and two comparisons (red and green hues), a task that should not induce a single-code / default strategy (see Grant & Spetch, 1994). Afterward, these pigeons ran a series of retention test trials. If the pigeons used the onset of the comparisons as the end of the to-be-timed interval, they would effectively be timing an empty interval with length "sample duration + retention interval". Hence, on test trials, we would expect preference for the "long" key to increase with the retention interval. On the other hand, if the sample offset was the "stop" time-marker, and because no default "long" response rule was acquired, we would expect preference for the "short" key to increase with the retention interval (choose-short effect), or indifference between the two comparisons (Sherburne et al., 1998). As a replication, a second group of pigeons learned the same task as in Study I, with three samples (2 s, 6 s and 18 s) and two comparisons (red and green hues). After the retention tests, the two groups of pigeons switched tasks to determine whether differences in training could result in differences in coding that would affect performance on the tasks.

Finally, to identify the stimulus used as the "start" time marker – houselight offset or sample onset – we ran two retention-interval tests without the sample, one with a houselight-illuminated ITI, and one with a dark ITI. In the first test, if the houselight was irrelevant and the sample was the main source of stimulus control, then, absent the sample, preference should not vary with the retention interval. The pigeons could show a bias for one comparison, but the bias should not change with the retention interval. Alternatively, if the pigeons learned to time the interval since houselight offset, preference for the "long" key should increase with retention interval. In the test with the dark ITI, without the two main sources of stimulus control, the pigeon was left without any cue about the appropriate choice. Hence, we expected indifference. Because the two no-sample tests differed only in the presence or absence of the houselight during the ITI, by contrasting the two test results we can assess the importance of the houselight on the animals' performance. The results from

the various tests may help to unravel the sources of temporal control in a widely used temporal discrimination task.

Method

Subjects

Eight pigeons (*Columba livia*) were maintained at approximately 80% of their free-feeding body weight, with water and grit freely available in their home cages. The pigeon room was maintained in a 13:11 hour light/dark cycle, with lights on at 08:00, and its temperature was maintained between 20-22 ℃. The experiment was conducted once a day, seven days a week, at approximately the same time of day for each pigeon. Three of the pigeons were experimentally naïve, and the remaining five had experience with timing tasks, either Fixed Interval schedules (Pinto, Fortes, Jozefowiez, & Machado, 2012), or matching-to-sample tasks with temporal samples (Carvalho & Machado, 2012).

Apparatus

Five LVE (Lehigh Valley Electronics) and a homemade chamber were used. The LVE chambers measured 34 x 35 x 31 cm (h x l x w). Three circular response keys, 2.5 cm in diameter, were arranged horizontally on the response panel, 9 cm apart, center to center. The bottom of each key was 22.5 cm above the wire mesh floor. Behind each key was a 12-stimulus IEE (Industrial Electronics Engineers) in-line projector. On the wall opposite the response panel, 30 cm above the floor, a 28-V, 0.1-A houselight provided general illumination. The food hopper was accessible through a 6-cm wide x 5-cm high opening on the response panel, centered horizontally, 8.5 cm above the floor. When the hopper was raised to provide grain to the pigeon, a 28-V, 0.04-A light illuminated its opening. The operant chamber was enclosed in an outer box equipped with an exhaust fan.

The homemade chamber measured 31 x 33 x 33 cm (h x l x w). Three circular response keys, 2.5 cm in diameter and 9 cm apart, (center to center) were arranged horizontally on the response panel. The bottom edge of each key was 21 cm above the wire mesh floor. For the presentation of stimuli, a 12-stimulus IEE (Industrial

Electronics Engineers) in-line projector was installed behind each key. A LVE food hopper was accessible through a 6-cm wide x 4.5-cm high opening that was centered horizontally on the response panel, 6.5 cm above the floor. When the hopper was raised, a 28-V, 0.04-A light illuminated its opening. On the wall opposite the response panel, 27.5 cm above the floor, a 28-V, 0.1-A houselight provided general illumination. The operant chamber was enclosed by a PVC sound attenuating cubicle (Med Associates, ENV-018V) equipped with an exhaust fan.

In this experiment, the side keys were illuminated with red or green hues and the central key was illuminated with a white hue. Personal computers using the ABET II software (Lafayette Instrument Company) controlled the experimental events and recorded the data.

Procedure

Training I. The animals were divided in two groups, which learned two symbolic matching-to-sample tasks in different orders, a task with two samples (2 s and 18 s) and a task with three samples (2 s, 6 s and 18 s). A trial began with the illumination of the center key with white light for 2, 6 or 18 s. At the end of the sample, the center keylight was turned off and the side keys were turned on, one with red and the other with green light. One comparison was correct following the 2-s sample, and the other comparison was correct following both 6-s and 18-s samples (the correct comparison for each sample was counterbalanced across pigeons). One peck at a comparison turned both keylights off. If the choice was correct, the pigeon had access to mixed grain and then the ITI started; if the choice was incorrect, the ITI started immediately. The ITI was illuminated with the houselight and lasted 30 s. Throughout the experiment, the houselight was illuminated *only* during the ITI.

When a response was incorrect, the trial repeated (correction procedure); following three consecutive incorrect responses, only the correct comparison was presented. Excluding correction trials, a session comprised 64 trials. For the two-sample group, there were 32 2-s sample trials and 32 18-s sample trials, and for the three-sample group, there were 32 2-s sample trials, 16 6-s sample trials, and 16 18-s sample trials. That is, the number of times each comparison was correct was the same for all birds. The location of the comparison stimuli varied randomly across trials,

with the constraint that each comparison occurred the same number of times on the left and right keys at the end of each session.

Training I continued for a minimum of 15 sessions and until the pigeons reached a criterion of at least 80% correct responses following each sample (excluding correction trials) for five consecutive sessions, or 25 sessions had elapsed, whichever occurred first. The birds began each session at approximately 80% of their free-feeding weight and the reinforcement duration, which was adjusted for each bird to minimize feeding outside the experimental session, varied from 3 to 5 s across animals.

Retention Test I. In this phase, a retention interval was introduced between sample offset and comparison onset. Table 1 summarizes the session structure for both groups. The retention interval could be 2.5-, 5-, 10- or 20-s long, and it was spent in darkness. A session comprised 80 trials. For the two-sample group, there were 48 regular training trials (24 of 2 s, 24 of 18 s) and 32 retention-interval test trials (16 following each sample, 4 for each retention interval duration). For the three-sample group there were 48 regular training trials (24 of 2 s, 12 of 6 s, 12 of 18 s) and 32 retention-interval test trials (16 following the 2-s sample, 4 per interval duration, and 8 following the 6-s and 18-s samples, 2 per interval duration).

A correct response was always reinforced (both during training and test trials), but the correction procedure was in effect only on regular trials. To minimize extra-session feeding, the reinforcement durations were readjusted and varied from 3 to 5 s across animals. This phase lasted five sessions.

Training II. Training II was the same as Training I with the exception that the tasks were reversed: The group that initially learned the two-sample discrimination now learned the three-sample discrimination and vice versa. Reinforcement durations were recalculated and varied from 2.5 to 7 s across animals.

Retention Test II. This phase was similar to Retention Test I, but with a different number of samples. That is, a pigeon that on Retention Test I ran the two-sample task, now ran the three-sample task, and vice versa. Reinforcement durations were recalculated and varied from 2.5 to 5 s across animals.

Retraining I. To re-establish a common, baseline performance before the no-sample tests, all pigeons returned to the three-sample discrimination task. The stability criterion remained the same as before.

Lit ITI, No-sample Retention Test. In this phase, no-sample trials were interspersed among regular training trials. In a no-sample trial, after the ITI elapsed the houselight was turned off and, following an interval, the comparison keys were presented. This interval, spent with all lights off, could be 0-s long (the comparison keys were presented immediately after the ITI), 2.5-, 5-, 10- or 20-s long, in which case it was identical to the retention interval used in previous phases.

Each session comprised 78 trials, 48 regular training trials (24 of 2 s, 12 of 6 s, 12 of 18 s) and 30 no-sample trials (six trials x five retention intervals). No-sample trials were never reinforced. Reinforcement durations on regular trials were recalculated and varied from 2.5 to 4 s across animals. Testing lasted 10 sessions.

Retraining II. This phase was exactly the same as Retraining I, with the sole exception that, as soon as the performance criterion was met for two sessions, the pigeon moved to the next phase.

Dark ITI, No-sample Retention Test. This test was the same as the No-sample Retention Test, with one exception: The ITI before no-sample trials was not illuminated. Therefore, before a no-sample trial began, the box was in darkness for the duration of the ITI, and continued dark for the duration of the retention interval in the no-sample trial, until the comparison keys were turned on. Reinforcement durations on regular trials varied from 2.5 to 5 s.

Results

Training (I and II) & Retraining (I and II). The birds learning the two-sample task first needed from 15 to 18 sessions ($\bar{x}=16$) to reach the criterion. Two of the four birds learning the three-sample task first needed 18 and 23 sessions, but the other two birds did not reach the criterion after 25 sessions. Although percent correct fluctuated across sessions, both birds performed significantly above chance: For each sample, the average over the last 5 session of percentage of correct was greater than 80%.

The birds needed from 15 to 22 sessions ($\bar{x}$ = 17) to reach criterion on the second task. Similarly, they needed from 15 to 23 sessions ($\bar{x}$ = 16) to complete Retraining I, and from 2 to 5 sessions ($\bar{x}$ = 3) to complete Retraining II.

Retention Tests (I and II). Figure 9 shows the results of the two retention tests, Test I on the left panel and Test II on the right panel. Each row of panels refers to the same pigeons, those who learned the two-sample task first (top) and those who learned the three-sample task first (bottom).
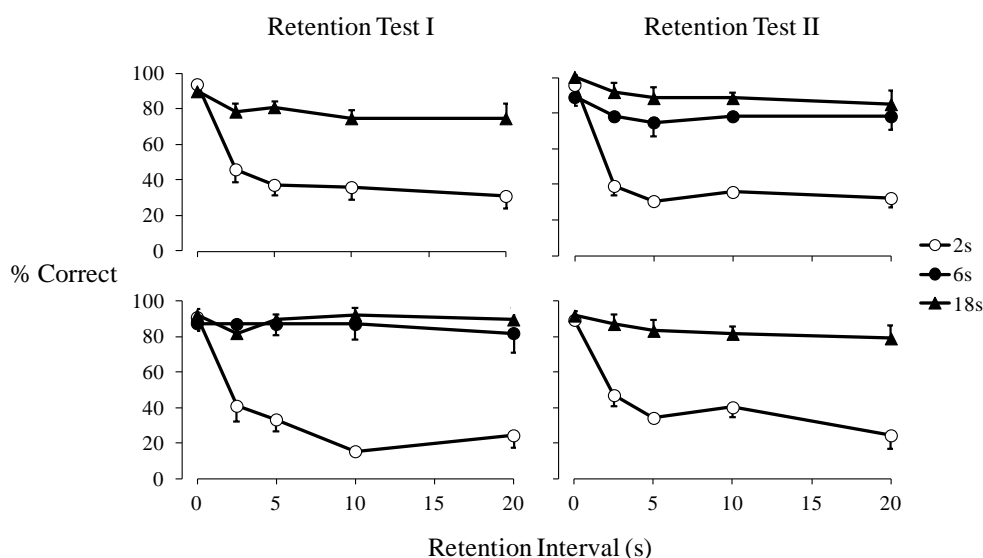
Figure 9. Mean (with SEM) percent correct to each sample duration as a function of retention interval duration. The panels on the left refer to Retention Test I and the panels on the right refer to Retention Test II. The panels on the top refer to the group of pigeons that first learned the task with two samples and then learned the task with three samples. The panels on the bottom refer to the group of pigeons that learned the tasks in reverse order.

Consider the results of the Retention Test I (left panels). Both groups showed the same pattern: As the retention interval increased, correct responses following 2-s samples decreased, while correct responses following the 6-s and 18-s samples remained high. For the two-sample group (top left panel), a two-way repeated-measures ANOVA with sample duration (two levels) and retention interval (five levels) as factors revealed a significant main effect of sample duration, $F(1, 3) = 18.8$,

$p = .023$, $\eta_G^2 = .72$, and of retention interval, $F(4, 12) = 26.60$, $p < .001$, $\eta_G^2 = .69$. The interaction also was significant, $F(4, 12) = 8.73$, $p = .002$, $\eta_G^2 = .47$, confirming that matching accuracy following the two samples was affected differently by the retention interval. Moreover, average percent correct following the 2-s samples on retention-interval trials was significantly below 50% (95% Confidence Interval = [32.5%-43.1%]).

For the three-sample group (bottom left panel), a 3x5 repeated-measures ANOVA showed a significant main effect of sample duration, $F(2, 6) = 27.93$, $p = .001$, $\eta_G^2 = .81$, retention interval, $F(4, 12) = 16.54$, $p < .001$, $\eta_G^2 = .43$, and their interaction, $F(8, 24) = 14.62$, $p < .001$, $\eta_G^2 = .57$, confirming that the effect of the retention interval was not the same for all samples. Similarly to the two-sample group, percent correct following 2-s samples on retention-interval trials was significantly below indifference (95% Confidence Interval = [22.9%-32.7%]).

Consider the Retention Test II (Figure 9, right panels). We found a similar pattern to Retention Test I: With the retention interval, performance following 2-s samples decreased below indifference, whereas performance following the 6-s and 18-s samples remained high. Therefore, the pattern of choices in Retention Test I was maintained in Retention Test II. For the group with three samples (top right panel), this reading was confirmed by a 3x5 repeated-measures ANOVA, which showed a significant effect of sample duration, $F(2, 6) = 21.01$, $p = .002$, $\eta_G^2 = .79$, retention interval, $F(4, 12) = 59.87$, $p < .001$, $\eta_G^2 = .53$, and their interaction, $F(8, 24) = 7.85$, $p < .001$, $\eta_G^2 = .49$. Percent correct following 2-s samples on retention-interval trials was significantly below 50% (95% Confidence Interval = [21.4%-31.1%]).

Regarding the group with two samples (bottom right panel), a 2x5 repeated-measures ANOVA revealed a significant main effect of sample duration, $F(1, 3) = 86.14$, $p = .003$, $\eta_G^2 = .81$, retention interval, $F(4, 12) = 14.59$, $p < .001$, $\eta_G^2 = .68$, and their interaction, $F(4, 12) = 9.75$, $p = .001$, $\eta_G^2 = .50$, once again confirming the differential effect of retention interval following the two samples. Percent correct following 2-s samples on retention-interval trials was significantly below indifference (95% Confidence Interval = [30.4%-40.9%]).

No-sample Retention Tests with Lit and Dark ITI. Although all pigeons went through the same baseline condition with three-samples before the no-sample tests, they had different histories. Hence, we first compared performance on the No-sample Test as a function of group. A repeated-measures ANOVA with one between-subjects factor (group) and one within-subjects factor (retention interval) showed no significant effect of group, $F(1, 6) = .026$, $p = .876$, $\eta_G^2 = .003$, or factor interaction, $F(4, 24) = .694$, $p = .603$, $\eta_G^2 = .03$, confirming that differences in learning history were not affecting performance on this task. Therefore, data from all animals were collapsed and analyzed together.
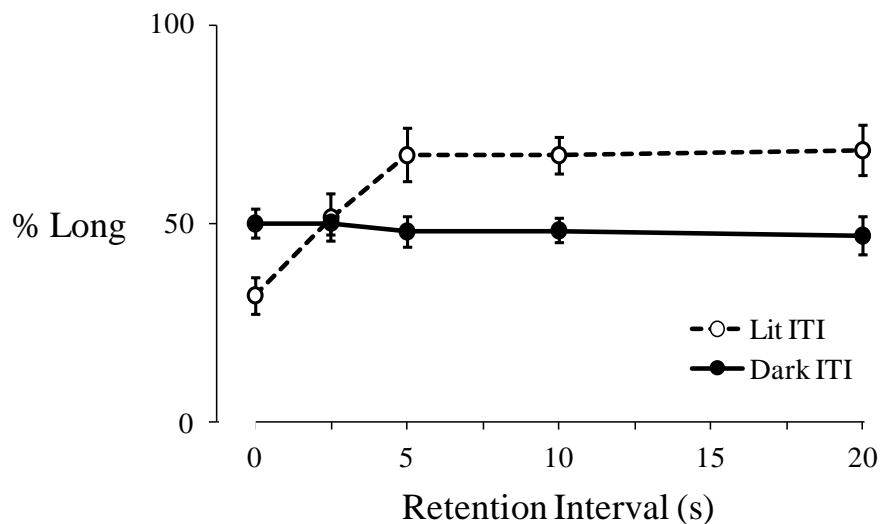
Figure 10. Mean (with SEM) percent of choices to the "long" key (associated with 6-s and 18-s samples) as a function of retention interval on the Lit ITI, No-sample Retention Test (empty circles) and on the Dark ITI, No-sample Retention Test (filled circles).

The empty circles in Figure 10 show the result of the Lit ITI, No-sample Test. On trials without a retention interval, the pigeons preferred the "short" key. A t-test confirmed that this preference was significantly different from 50%, $t(7) = 3.99$, $p = .005$, $d_z = 1.41$, 95% CI for d [0.39, 2.39]. On trials with retention intervals, the preference shifted towards the "long" key. A one-way repeated-measures ANOVA revealed a significant main effect of the retention interval, $F(4, 28) = 18.93$, $p = .001$,

η_G^2 = .46. By the 20-s retention interval, the preference for the "long" key was significantly above chance: $t(7) = 2.85$, $p = .025$, $d_z = 1.01$, 95% CI [0.12, 1.85].

The filled circles show the results from the Dark ITI, No-sample Retention Test. The pigeons were indifferent between the two comparison keys, and preference did not vary with the retention intervals. A one-way repeated-measures ANOVA showed no significant effect of retention interval, $F(4, 28) = .784$, $p = .545$, η_G^2 = .02. Additionally, choices were not significantly different from chance (95% Confidence Interval = [46.4%-50.5%]).

Discussion

In the present paper, we aimed to identify the "start" and "stop" time markers in a temporal discrimination task by using two types of tests, retention interval and no-sample tests. The retention interval tests examined which stimulus signalled the end of the duration that controlled the pigeons' choices. The birds were first divided in two groups, one that should have learned a single-code / default strategy to map three samples onto two comparisons, and one that should not (two samples and two comparisons). The three-sample group replicated Study I and thereby checked the reliability of its findings. The results were similar, a strong preference for the "long" key (Figure 9, lower left panel). The other group tested the key idea: In the absence of a single-code / default strategy, a strong preference for the "long" key would suggest the use of the comparisons onset as the "stop" time marker. Introducing the retention intervals would make the effectively-timed interval longer, which would yield a strong preference for the "long" key. The results (Figure 9, upper left panel) were consistent with the hypothesis.

Subsequently, the groups learned the other discrimination task and again ran retention interval tests. The preference functions in Retention Test II (Figure 9, right panels) were consistent with the preference functions in Retention Test I (Figure 9, left panels). Both showed a strong preference for the "long" key. The fact that the preferences remained the same on both sets of retention tests suggests that the order in which the tasks were learned did not affect the coding strategy.

The no-sample tests assessed which stimulus functioned as the "start" time marker. By removing the sample stimulus, we aimed to assess directly the role of the houselight offset as a time marker. If sample onset triggered timing, by removing the sample, timing should not take place, which should result in random responding. But if the houselight offset was the "start" time marker, we would expect choices to correlate with the retention interval: the longer the retention interval, the longer the interval since the ITI offset, and therefore the stronger the preference for the "long" key. The results from the Lit ITI, No-sample Retention Test (Figure 10, empty circles) were consistent with the latter prediction. Choices immediately following the ITI were mostly to the "short" key, but, as the retention interval increased, preference shifted towards the "long" key.

In the second no-sample test, the houselight remained off during the ITI. This manipulation complemented the Lit ITI, No-sample test because, by comparing the two no-sample tests, we could assess how behavior was affected by the removal of the houselight. The results showed random choice across all retention intervals (Figure 10, filled circles), a result distinctly different from the Lit ITI, No-sample test. The change in performance occasioned by removing the houselight during the no-sample tests shows clearly that the houselight influenced choice.

Taken together, the results of the no-sample tests lend further support to the hypothesis that the pigeons used the houselight offset as the "start" time marker. However, the data also suggest that the houselight was not the sole time marker. If choices were made exclusively on the basis of how much time had elapsed since the offset of the houselight, on no-sample trials choices would approach the values obtained during regular trials with comparable samples. For instance, a no-sample trial with a 2.5-s retention interval should result in a percentage of "short" responses similar to that found on regular trials following 2-s samples. However, that was not the case. The pigeons chose the "short" key 94% of the time on 2-s regular trials, but that preference fell to 49% on no-sample trials with a 2.5-s retention interval. The same was true for longer values: A no-sample trial with a 20-s retention interval should result in a percentage of "long" key choices similar to that found on regular trials following 18-s samples. Again, that was not the case – the pigeons chose the

"long" key 98% of the time on 18-s regular trials, but that preference fell to 68% on no-sample trials with a 20-s retention interval.

To summarize, even though the evidence supports the notion that the houselight offset functioned as a time marker, the full data set is not consistent with the houselight as the only stimulus controlling choice. Perhaps both the houselight and the center keylight (sample) exerted control over responding. Performance worsened when the sample stimulus was removed *and* also when the ITI was darkened (by this point the pigeons were responding randomly). The fact that the removal of each of these stimuli had an effect on performance suggests that both controlled choice. This joint control can be seen in Figure 11, where all average data were replotted as a function of time since the houselight offset. Each line represents a sample duration. If houselight offset were the only "start" time-marker (i.e., the sample stimulus was irrelevant), the curves should overlap. The fact that they do not overlap shows that the sample also controlled choice.
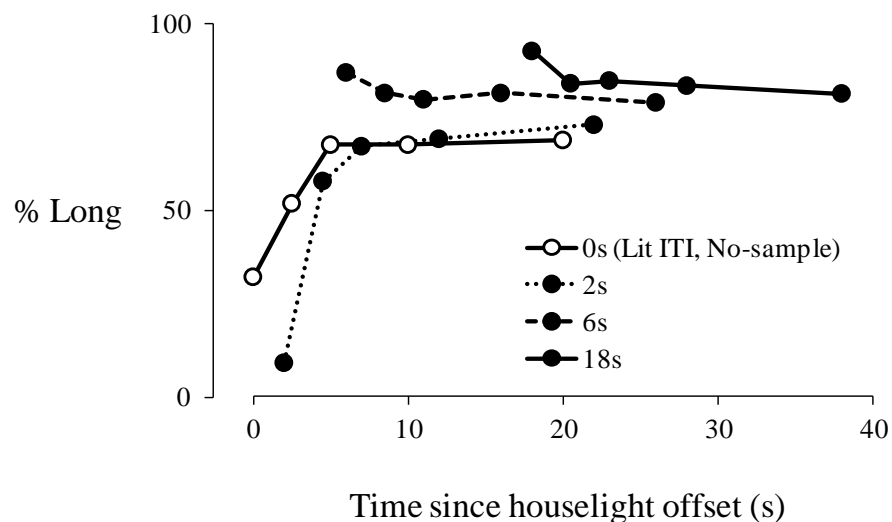

Figure 11. Mean percent of choices to the "long" key (associated with 6-s and 18-s samples) as a function of time since houselight offset. Each line refers to a sample duration. The filled circles represent the collapsed data from Retention Test I and II and the empty circles represent the data from the Lit ITI, No-sample Retention Test.

Not all studies with illuminated ITIs and dark retention intervals have found a similar choose-long effect. For instance, Sherburne et al. (1998) found that, on a retention test, matching accuracy did not differ between the two trained samples (2 s and 10 s). The reason for the different results may be related to the fact that in Sherburne et al. (1998) the 10-s ITI was considerably shorter than the 30-s ITI used in the present experiment. A shorter ITI may have made the houselight less likely to control choice.

In another study, Kelly and Spetch (2000) obtained a bias towards the "short" key (choose-short effect) in a task that had a long ITI (45 s), but used a 5-s retention interval during training. In another study that employed a retention interval during training (but with variable duration), a choose-long effect, similar to the one found in this paper, was obtained (Dorrance, Kaiser, & Zentall, 2000). It is not entirely clear how task differences lead to distinct retention-test functions, but the variety of the results suggests that a houselight-illuminated ITI per se does not make the houselight a time marker.

Another variable that influences the coding strategy adopted in a matching-to-sample task is the sample-comparison mapping. There has been some evidence that many-to-one mappings eliminate the choose-short effect (Grant & Spetch, 1993; Santi, Bridson, & Ducharme, 1993). However, in those studies, the choose-short effect was not replaced by the choose-long effect found in the present work, which reinforces the hypothesis that the preference for the "long" comparison we found was not due to the sample-comparison mapping, but to the influence of the ITI houselight.

The present study shows that, in a temporal matching-to-sample task, other stimuli besides the sample may control responding. This result supplements previous research that has shown that more than one stimulus or event may concurrently control responding (e.g., Cheng, Spetch, & Miceli, 1996; Roberts & Mitchell, 1994), and that animals are able to simultaneously time more than one interval (e.g., Kirkpatrick & Church, 2000; Leak & Gibbon, 1995; Meck & Church, 1984). To understand how and what animals learn, even in simple temporal discrimination tasks, it is important to identify the multiple sources of stimulus control present in the task. In addition, we also need to identify the procedural features that occasion joint

stimulus control and the determinants of the relative degree of stimulus control achieved by each source.

JOINT STIMULUS CONTROL IN A TEMPORAL DISCRIMINATION TASK[4]

To investigate the nature of joint stimulus control in a timing task, pigeons learned a matching-to-sample task with three durations as samples (2 s, 6 s and 18 s of keylight) and two colors as comparisons (red and green hues). A 30-s intertrial interval (ITI), illuminated with a houselight, separated the trials. Because both the houselight offset and the sample keylight onset could function as time markers, two tests were run to assess whether and how these stimuli controlled choice. In the no-sample test, the keylight was not presented; in the dark-ITI test the houselight was not illuminated. Results suggest that both houselight offset and keylight onset controlled choice. Moreover, the more one pigeon relied on one of these stimuli, the less it relied on the other. We present a quantitative model of stimulus competition that estimates the relative degree of control of each stimulus over responding.

Not all stimuli are equally important to understand how an animal adapts to its environment. Stimuli that correlate with important events (e.g., food, predators, or partners), tend to be attended; those that do not, tend to be ignored. Moreover, not all features that define a stimulus (e.g., size, color, or shape of a visual stimulus; loudness, frequency, or location of an acoustic stimulus; nature and intensity of a smell, etc.), may be equally informative in a given situation. Naturally, then, many animals can learn to respond selectively to the critical stimulus features in their surroundings (e.g., Sutherland & Mackintosh, 1971).

Researchers have studied attention to stimuli in a variety of ways. In one of them, they have used stimuli with multiple elements, compound stimuli, and attempted to identify the stimulus features an animal attends to. In a classic study, Reynolds (1961) trained two pigeons with a compound stimulus, a key displaying a white triangle on a red background, and then tested them with the elements of the compound separated. He found that, whereas one pigeon pecked mostly to the red hue, another pecked mostly to the triangle. The two birds seemed to have attended to different elements of the compound. Other studies have used compounds with elements from the same modality (e.g., visual stimuli: Leith & Maki, 1975; Maki & Leith, 1973) or from different modalities (e.g., visual + auditory stimuli: Blough, 1969; Kraemer & Roberts, 1985; spatial + visual stimuli: Kraemer, Mazmanian, & Roberts, 1987).

When a task has two redundant cues such that relying on only one of them is sufficient for correct performance, animals may nevertheless attend to both. Roberts and Mitchell (1994) trained pigeons to discriminate between 2 flashes of light (lasting, in total, 2 s) and 8 flashes of light (lasting, in total, 8 s). The pigeons could have learned to respond based on the number of flashes (2 vs. 8) or the cumulative duration of the stimuli (2 s vs. 8 s). Subsequently, the authors ran two types of test trials, one in which the number of flashes remained constant at 4 flashes while cumulative duration varied from 2 to 8 s, and another in which cumulative duration remained constant at 4 s while the number of flashes varied between 2 and 8. Results showed that the pigeons attended to both number and duration (see also Meck & Church, 1983).

In another study, Cheng, Spetch and Miceli (1996) ran a task where pigeons could attend to duration or location. A white rectangle moved horizontally in a touch-sensitive screen at constant speed. The first peck on the rectangle after 10 s had elapsed granted access to food. Because the rectangle moved at constant speed, its location also signaled reinforcement. Therefore, the pigeons could have learned either to wait 10 s before pecking, or to peck when the rectangle reached a specific location of the screen. On test trials, the authors varied the speed of the rectangle and found that both time and location influenced responding; the two dimensions shared stimulus control (see also Sutton & Roberts, 1998).

Even when tasks are not designed to study joint stimulus control, they may provide alternate sources of information an animal can use. In interval timing tasks, for example, more than one event may mark the to-be-timed interval. In studies I and II, we trained pigeons in a matching-to-sample task in which a center key was illuminated with a white hue for 2, 6 or 18 s (sample) and then a choice was given between a green and a red side key (comparisons). The red key was correct if the light was on for 2 s, and the green key was correct if the light was on for 6 or 18 s. During the 30-s intertrial interval (ITI), a houselight illuminated the whole box. Even though the task was designed assuming that the pigeons would learn the task by timing the duration of the sample stimulus, the interval from the houselight offset to the comparisons onset was also a valid time marker. A series of test trials with retention intervals, with and without the sample, and with dark ITIs suggested that the pigeons attended to both the duration of the white keylight and the interval since the houselight was turned off. Evidence for joint control in temporal matching-to-sample tasks is particularly important because researchers often assume that their subjects are timing *only* the sample stimulus. If other stimuli control performance in these and similar tasks, it is important to identify them and to understand how two or more sources of stimulus control relate to one another. Without this knowledge our ability to predict and understand behavior will remain limited.

In the present experiment we used the 3-sample, 2-comparison task described above to examine the nature of joint stimulus control. First, to contrast the influence of the two putative time markers, the sample keylight onset and the houselight offset, after the pigeons had learned the task we ran two types of tests trials. On each test

trial, one of the time markers was removed but the other remained unchanged. In the no-sample test, the sample keylight was removed and the choice keys were made available immediately after the houselight offset that ended the ITI. If timing was triggered by the houselight offset, the pigeons should prefer the comparison associated with the shortest sample: Without the sample, the timed interval would be shorter than in training and, by temporal generalization, the pigeons should be biased toward the short comparison. If, on the other hand, animals were timing the duration of the keylight and timing was triggered by the sample onset, random responding would be expected because no sample was presented, so no timing would have occurred. In the dark-ITI test, the houselight was never turned on, so the ITI was spent in darkness. If animals were timing the duration elapsed since the presentation of the houselight and timing was triggered by the houselight offset, because timing would continue since the previous trial, the pigeons would prefer the "long" comparison, making most errors following the 2-s samples. In contrast, if timing was triggered by the sample onset, performance in dark-ITI tests should not be disrupted, or at least the different samples should not be affected differentially; responding should remain similar to the training trials. Table 1 summarizes the predictions. In conjunction, the no-sample and dark-ITI tests allowed us to estimate the influence of keylight onset and houselight offset on choice.

Table 1. Predicted preferences in no-sample and dark-ITI tests, as a function of stimulus controlling responding.

Control by	Test	
	No sample	Dark ITI
Houselight offset	"short" key	"long" key
Sample onset	Indifference	Baseline performance

Second, we evaluated the possibility that both time-markers affected responding in a competitive way: The more an animal relies on one stimulus, the less it may rely on the other (e.g., Blough, 1969; Kirkpatrick-Steger & Wasserman, 1996; Vyazovska, Teng, & Wasserman, 2014; see also Thomas, 1970).

As an animal is exposed to a task, its reliance on alternate cues may change. To prevent such changes from biasing our measures of stimulus control, each test session included both types of test trials, so that the influence of each stimulus could be assessed concurrently. Finally, by taking together the individual results from the two types of tests, we put forward a quantitative model that estimated the relative influence of each time marker on responding.

Method

Subjects

Ten pigeons (*Columba livia*), were maintained at approximately 80% of their free-feeding body weight. The animals were kept in individual home cages, where water and grit were freely available. The pigeon room, kept between 20-22 ℃, was maintained in a 13:11 hour light/dark cycle, with lights on at 08:00. The experiment was conducted once a day, at approximately the same time of day for each pigeon, seven days a week. Seven pigeons had no experience with matching-to-sample tasks.

Apparatus

Four identical LVE operant chambers, measuring 34 x 35 x 31 cm (h x l x w), were used. On the response panel, three circular response keys, 2.5 cm in diameter, and 9 cm apart, center to center, were arranged horizontally. The bottom of each key was 22.5 cm above the wire mesh floor. Each key was equipped with a 12-stimulus IEE (Industrial Electronics Engineers) in-line projector. The food hopper was accessible through a 6-cm wide x 5-cm high opening, centered horizontally on the response panel, 8.5 cm above the floor. When the hopper was raised, a 28-V, 0.04-A light illuminated its opening. On the wall opposite the response panel, 30 cm above the floor, a 28-V, 0.1-A houselight provided general illumination. The operant

chamber was enclosed in an outer box equipped with an exhaust fan that circulated air through the chamber and masked outside noises.

In this experiment, red or green hues were presented on the side keys and a white hue was presented on the central key. Personal computers running the ABET II software (Lafayette Instrument Company) controlled the experimental events and recorded the data.

Procedure

Training. The pigeons learned a symbolic matching-to-sample task. A session started with a 30-s intertrial interval (ITI), during which the houselight was turned on. Afterwards, the houselight was turned off and the center key was turned on with a white hue for 2, 6 or 18 s (sample). At the end of the sample, the center keylight was turned off and the two side keys were illuminated, one with a red and the other with a green hue (comparisons). One comparison was correct following the 2-s sample, and the other comparison was correct following the 6-s and 18-s samples. The correct comparison for each sample was counterbalanced across pigeons. After a peck, the two comparison keylights were turned off. If the response was correct, reinforcement was delivered and then the ITI started; if the response was incorrect, the ITI started immediately. A correction procedure was in effect: Following an incorrect response, the trial was repeated; after three consecutive incorrect responses, only the correct comparison was presented.

The birds began each session at approximately 80% of their free-feeding weight. To minimize feeding outside the experimental session, reinforcement duration was adjusted individually, and varied from 1.5 s to 6 s across animals.

Each session comprised 64 trials (excluding correction trials), 32 2-s sample trials, 16 6-s sample trials, and 16 18-s sample trials. These values meant that each comparison was the correct choice the same number of times in each session. Across trials, the location of the comparisons varied pseudo-randomly with the constraint that each comparison was presented the same number of times on each side key. Training lasted a minimum of 15 sessions, and it continued until the pigeon met a criterion of at least 80% correct responses to each sample in a session (excluding correction

trials), for five consecutive sessions, or until a maximum of 30 sessions was completed.

Testing. Each session included three types of trials, regular training trials, no-sample test trials, and dark-ITI test trials. The training trials remained exactly as during the Training phase. On the no-sample trials, the center key was not illuminated and the comparisons followed the ITI immediately. On the dark-ITI trials, all procedural details remained as on the training trials except that the houselight was not turned on during the ITI.

The three types of trials were arranged in blocks. A Training block included 16 training trials (8 x 2 s, 4 x 6 s, 4 x 18 s). A no-sample block included 8 no-sample trials randomly interspersed among 8 training trials (4 x 2 s, 2 x 6 s, 2 x 18 s), for a total of 16 trials. No-sample trials were never reinforced. Lastly, a dark-ITI block included 16 training trials (8 x 2 s, 4 x 6 s, 4 x 18 s) in which the ITI preceding the sample was spent in darkness. On dark-ITI trials, correct responses were reinforced.

Every test session comprised four 16-trial blocks, for a total of 64 trials. The first and third blocks were always training blocks; the second and fourth blocks were either no-sample and dark-ITI blocks, respectively, or dark-ITI and no-sample blocks, respectively. The order of the no-sample and dark-ITI blocks alternated across days. On the first day, their order was counterbalanced across pigeons with half experiencing one order, and the other half the other order. Testing lasted 10 sessions.

Results

Training. The birds completed the Training phase in 26 sessions on the average (range: 15 - 30). Although six birds failed to reached the criterion of 80% correct responses to each sample for five consecutive days, they clearly learned the discrimination: For those birds, percent correct following each sample on the last five sessions averaged 78% or above. Overall, during the last 5 sessions of training, matching accuracy averaged 90% on 2-s trials, 85% on 6-s trials, and 95% on 18-s trials.

Testing. During the test phase, performance on training trials remained similar to the training phase: matching accuracy over the 10 test sessions averaged

94% following 2-s samples, 85% following 6-s samples and 97% following 18-s samples. Therefore, testing did not disrupt baseline performance.

On the no-sample test trials, nine pigeons preferred the "short" comparison ($\bar{x}$ = 73%; range: 45% - 93%). A t-test confirmed that this preference differed significantly from 50% chance, $t(9) = 4.56$, $p = .001$, Cohen's $d_z = 1.44$, 95% CI for d [0.52, 2.33]. The preference for the "short" key on no-sample tests replicates our findings from the previous studies, and it suggests that the pigeons used the houselight as a time marker (cf. Table 1). However, if they were exclusively timing the interval since the houselight offset, an even stronger preference for the "short" comparison would be expected.

Figure 12 shows the results from training trials and dark-ITI trials. Removing the houselight illumination during the ITI seemed to reduce accuracy only following the 2-s samples. This visual impression was confirmed by statistical analyses: A paired-samples t-test revealed that when the ITI was darkened correct choices following the 2-s samples decreased significantly ($t(9) = 5.19$, $p = .001$, $d_z = 1.64$, 95% CI [0.65, 2.59], but for the other samples, the changes were not significant (6-s samples, $t(9) = 1.70$, $p = .124$, $d_z = 0.54$, 95% CI [-0.14, 1.19]; 18-s samples, $t(9) = 2.00$, $p = .077$, $d_z = 0.63$, 95% CI [-0.07, 1.30]). In other words, the percentage of "long" responses increased when the ITI was not illuminated, a result also consistent with the use of the houselight as a time marker (cf. Table 1). However, if the animals were estimating exclusively the interval since the (last) houselight offset, the decrease in percent correct following the 2-s samples should have been more pronounced. The data suggest that the sample duration also influenced responding.
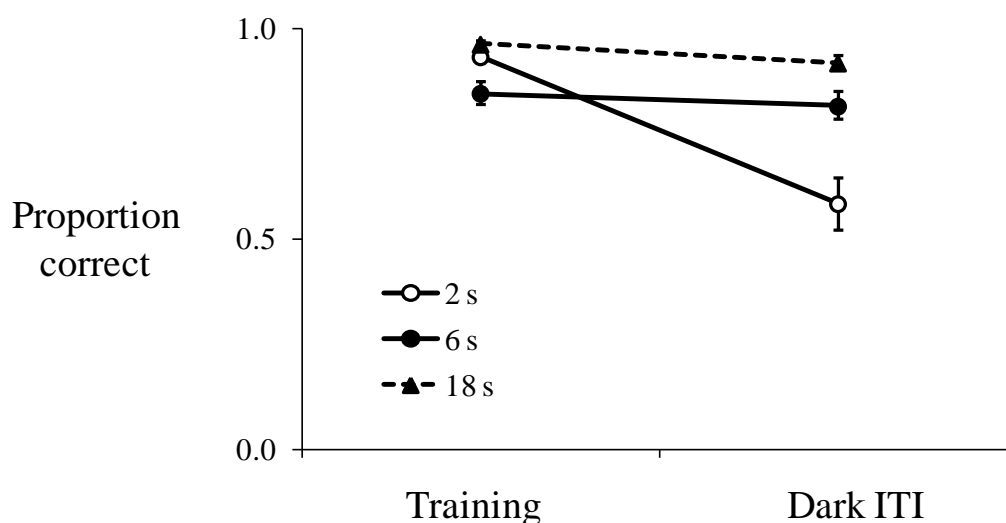
Figure 12. Mean (with SEM) proportion correct to each sample duration during testing, in training trials (left side) and in dark-ITI test trials (right side).

Discussion

If both the houselight offset and the center keylight (sample) onset controlled choice, then we may ask about the nature of their joint control. In what follows we assume two premises, shared with classic selective-attention models (Lovejoy, 1968; Mackintosh, 1965; Sutherland & Mackintosh, 1971; Zeaman & House, 1963), that a) there is competition between the two stimuli such that the more a pigeon relies on one, the less it relies on the other, and b) on a given trial the pigeon relies on only one of the two stimuli. By contrasting performance on the no-sample and dark-ITI tests, it is possible to estimate, for each individual subject, the relative effect of each stimulus on responding.

We base our analysis on two assumptions, one for each test type. First, on no-sample trials the percentage of choices of the "short" key correlates positively with the influence of the houselight on choice: across pigeons, the higher the relative influence of the houselight, the higher the preference for the "short" key. Second, on dark-ITI trials, the percentage of correct responses following 2-s samples correlates positively with the influence of the center keylight on choice: across pigeons, the

higher the relative influence of the keylight, the higher the percentage of correct responses following 2-s samples.

Consider then a pigeon that on each trial attends to the houselight offset with probability p, and to the center keylight (sample) onset with probability 1-p. These two possibilities combine with the two trial types to yield the four cases displayed in Table 2. On a no-sample trial, if the pigeon attends to the houselight offset it will time a very short interval and therefore, because of stimulus generalization, it will choose the "short" key; if it attends to the sample then, because there was no sample, it will choose the two comparisons randomly and P("short")=0.5. Alternatively, if the pigeon remembered the last sample presented (one or more trials ago), since the proportion of times each comparison was correct was equated, on average responding would also approach indifference. Therefore, the probability of choosing the "short" key in a no-sample trial is given by

$$P(\text{"short"}| \text{ No-sample trial}) = p\times1 + (1\text{-}p)\times0.5 = (1+p)\times0.5 \tag{1}$$

On a Dark-ITI trial with a 2-s sample, if the pigeon attends to the houselight offset then, regardless of sample duration, it will choose "long" because the last time the houselight was turned off occurred one or more trials ago and therefore the timed interval is longer than the longest training sample; however, if the pigeon attends to the sample then, on the 2-s sample trials it will choose "short" with the same probability of a regular training trial, α, say. Thus, the probability of choosing the "short" key in a 2-s dark-ITI trial is given by

$$P(\text{"short"}| \text{ Dark-ITI, 2-s trial}) = p\times0 + (1\text{-}p)\times\alpha = (1\text{-}p)\times\alpha, \tag{2}$$

These assumptions imply that the probability of choosing the "short" comparison correctly given a Dark-ITI trial, P("short"| Dark ITI, 2-s trial), is a linear function of the probability of choosing the "short" comparison given a No-sample trial, P("short"| No-sample trial). In fact, from Equations 1 and 2 it follows that

$$P(\text{"short"}| \text{ Dark-ITI, 2-s trial}) = 2\times\alpha - 2\times\alpha\times P(\text{"short"}| \text{ No-sample trial}), \tag{3}$$

or, more concisely, $Y = 2\alpha - 2\alpha X$. The two probabilities should be negatively correlated. Moreover, because $0 \leq p \leq 1$, $0.5 \leq X \leq 1$, and $0 \leq Y \leq \alpha$, when α is close to 1 (i.e., the pigeons learned to choose "short" on 2-s sample trials), these inequalities mean that the the Y and X variables have different ranges and, more specifically, that the (X, Y) data points will fall on the right half side of the X-Y graph.

Table 2. Probability of choosing "short" according to the time marker used and the trial type. p is the probability of using the houselight offset as the time marker and α refers to training performance.

Time marker	Test	
	No sample	Dark ITI, 2-s sample
Houselight offset (p)	P("short") = 1	P("short") = 0
Sample onset (1-p)	P("short") = 0.5	P("short") = α

Figure 13 compares the predictions of Equation 3 against the data, with parameter α set to .94, the average across pigeons of proportion "short" on 2-s trials. With one exception, discussed below (see unfilled circle), the solid line describes well the obtained trend, in particular the negative correlation between performance on the two trial types. Concerning the trend of the data points, a Pearson's correlation was marginally significant, $r = -.64$, $p = .06$. To evaluate the adjustment of the model, instead of using a common value for α, we used each pigeon's proportion correct on the 2-s sample trials, and found a significant correlation ($r = .67$; $p = .05$). Moreover, as the model predicted, the data points fell mostly on the right half of the unit square.
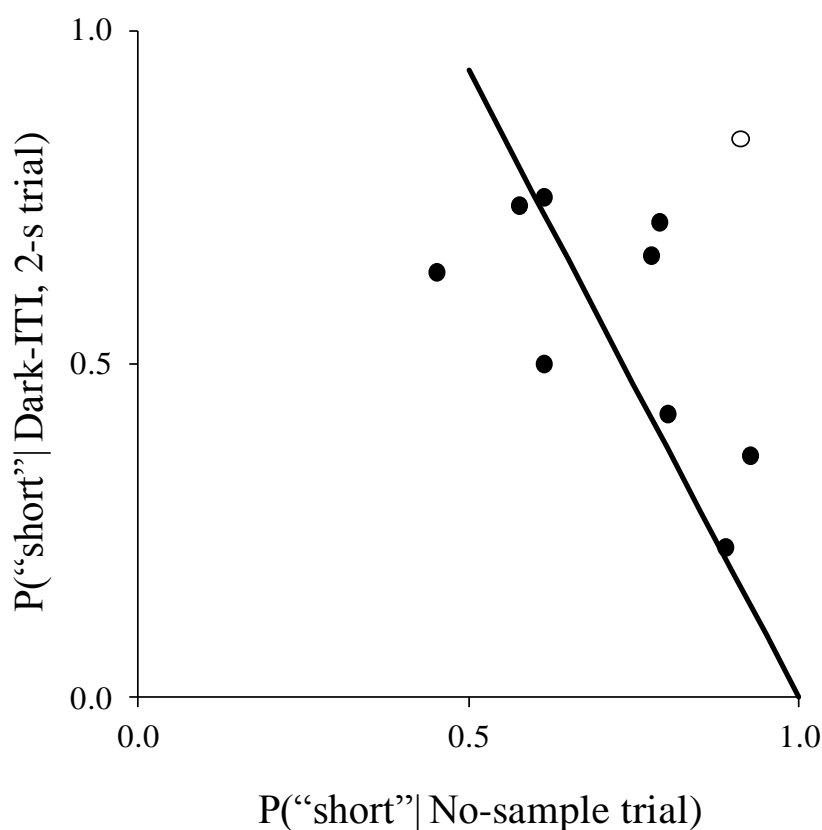
Figure 13. Proportion of choices to the "short" key in no-sample test trials (x axis) plotted against proportion of correct responses following 2-s samples in dark-ITI test trials (y axis). The line is the prediction of the model (with α = .94). Each dot refers to the performance of one pigeon. The white dot identifies pigeon PG29.

Pigeon PG29 was the exception. It showed high values on both axes, a result inconsistent with the assumed competition or trade-off between the two time markers. PG29 was the only pigeon whose matching accuracy following 2-s samples on dark-ITI trials remained above 80%, the learning criterion. This result suggests a strong reliance on sample onset as a time marker. However, the .91 proportion of "short" responses on the no-sample trials suggests a strong reliance on the houselight offset as a time marker. Hence, it seems that this pigeon may have learned to use both stimuli simultaneously as time markers.

To summarize, the results were consistent with the hypothesis that animals were attending to two different stimuli, and timing two intervals: the duration of the white keylight and the time elapsed since the offset of the houselight that ended the

ITI. This result confirms that, in a timing task, two temporal stimuli can jointly control responding. Among the 10 birds, one bird appeared to rely on the two stimuli simultaneously. For the remaining animals, a simple model based on the premise that there is a trade-off in the use of the two stimuli provided a reasonable approximation to the data.

Not all stimuli in the environment acquire control over responding, and if one assumes that it is more taxing to attend to two stimuli (or two aspects of a stimulus) than to a single stimulus, why would an animal attend to two different stimuli? It is possible that the benefit of having more than one source of information may compensate the additional cost. Moreover, the way a task is set up may facilitate the acquisition of control by more than one stimulus. In the present study, the finding that the houselight also was attended to may be explained by its salience: the houselight is a strong light that illuminates the whole box. Johnson and Cumming (1968) manipulated the intensity of one element in a compound visual stimuli and found that the higher the intensity of that element during training, the stronger its control over responding (see also Gaitan and Wixted (2000) for a detection theory based on stimulus salience).

When joint control does occur, it is likely to differ across subject; that is, the degree of competition between stimuli may vary and different stimulus control topographies may develop (Mackintosh, 1975; McIlvane & Dube, 2003). At the light of the model, differences in performance reveal the magnitude of the differences in the use of the available information. Thus, the estimated individual values for p, the proportion of trials on which the pigeons attended to the houselight, ranged from .09 to .77, suggesting that whereas some birds rarely attended to the houselight, others may have been influenced mostly by that stimulus. Similarly, Reynolds (1961) found that, when trained with a white triangle in a red background, one pigeon attended to the red hue and the other to the triangle. In a separate task, a side lamp was illuminated with either green or yellow (which differed in intensity, the yellow being brighter than the green), to indicate which element of the compound stimulus would provide reinforcement. Neither bird attended to the color of the lamp, and responded based on the intensity of the lamp illumination: when a red side lamp (as bright as the green) was used, pigeons responded as if a green lamp was presented; when a white

lamp (brighter than the yellow) was used, pigeons responded as if a yellow lamp was presented; finally, when a dim yellow light was used, pigeons responded as if a green lamp was presented.

When a task includes correlated stimuli, we may ask not only which stimuli will be used to solve the task but also how they interact. For instance, a conditioned response or the rate of operant responding may be higher in the presence of a combination of stimuli (each conditioned / trained separately) than in the presence of each stimulus individually (summation effect: Weiss, 1964; Wolf, 1963). On the other hand, it has been found that matching accuracy is lower when a compound sample is presented than when a single-element sample is presented (element superiority effect: Maki & Leith, 1973). That is, combining the elements seems to strengthen their individual effects in the first case, but to weaken them in the second case.

In conclusion, the present results stress the importance of taking into account, both when designing procedures and interpreting results, the possibility that a task may not be learned in the way expected by the experimenter (for example, our "matching-to-sample" task may predominantly be a "timing-since-ITI" task for some birds), and what is learned may vary significantly between subjects.

- 62 -

STUDY IV

CODING IN PIGEONS: FURTHER TESTS OF SAMPLE-COMPARISON MAPPING

Using the same base task as the previous studies, the sample-comparison mapping was changed so that the shortest and longest samples shared a comparison. This was done to assess the effect different mapping may have on what is learned, and to compare with a similar procedure used by Singer, Klein and Zentall (2006). Acquisition was slower than in our previous experiments, which points to the samples not being treated equally in learning. In a retention test, we did not replicate Singer et al. (2006)'s results, with animals seemingly adding sample and retention interval durations, a result consistent with the use of the houselight that signals the intertrial interval as a time marker. The difference between studies may be the salience of the stimuli used. In a no-sample test, a preference for the shared comparison was found. Taken together with the similar no-sample tests from previous studies, this result suggests that, when no sample is presented, animals prefer the comparison associated with the shortest sample.

To study the flexibility of the response rules that animals may use, Singer, Klein and Zentall (2006, Experiment 2) trained pigeons in a matching-to-sample task with three samples (white houselight illuminated for 2, 8, or 32 s) and two comparisons (red and green hues projected on the lateral keys). One comparison was correct following the 8-s sample, and the other comparison was correct following 2- and 32-s samples. To learn this task, three sample-specific response rules could be established, one for each sample (e.g., "If 2s, choose red", "If 8s, choose green", and "If 32s, choose red"). However, this task also allowed an alternative coding strategy, using only two response rules: "If 8s, choose green" and "If not 8s, choose red". This coding strategy is known as single-code / default, in the sense that a specific code is established for a single sample and there is a response by default if that one sample is not presented (Clement & Zentall, 2000; Colwill, 1984; Grant, 1991; Sherburne & Zentall, 1993; Weaver, Dorrance, & Zentall, 1999; Wilkie, 1978; Wilson & Boakes, 1985).

Singer et al. (2006) were interested in seeing whether pigeons would adopt the single-code / default strategy, a more economical solution to the task. To do so, after training, a retention interval was introduced between sample and comparisons. The reasoning was that, as the retention interval increases, the sample information is more likely to be lost, and therefore if the animal was responding according to the single-code / default rules, there should be a preponderance of choices to the "default" key (in the example in the previous paragraph, the red comparison). Therefore, accuracy should decrease with the retention interval following 8-s samples, but it should remain high following 2- and 32-s samples. That was the pattern found by Singer et al. (2006). However, this test was not conclusive, as the establishment of three sample-specific rules could have yielded similar results (for a discussion of this issue, see the introduction of Study I, p. 10). In an attempt to clarify the coding strategy, in Study I a task similar to Singer et al. (2006) was used, with three samples and two comparisons. However, the sample-comparison mapping was not the same: the sample that was associated with a specific comparison was the shortest duration.

One of the goals of the present study was to test the effect described by Singer et al. (2006) by using the same sample-comparison mapping, but maintaining the parameters used in studies I, II and III. That is to say, the sample stimulus was a white

hue on the center key (instead of a filtered houselight), the sample durations were 2, 6 and 18 s (instead of with 2, 8 and 32 s), and the intertrial interval was 30-s long (instead of 15 s). Since the parameters were maintained in regards to other tasks used in this dissertation, this experiment also invites a comparison with those other studies, to assess how a difference in mapping could affect performance, both in acquisition and testing.

When sample information becomes unavailable during a retention interval, it is assumed that the animal behaves as if no sample was presented. To test this assumption, a no-sample test was run: a trial started immediately with the presentation of the comparison stimuli. We were also interested in comparing performance in this test with similar no-sample tests ran in previous experiments of this dissertation.

Method

Subjects

Four pigeons (*Columba livia*) were maintained at approximately 80% of their free-feeding body weight. The colony room was illuminated in a 13:11 hour light/dark cycle, with lights on at 08:00, and it was kept between 20-22 ºC. Each animal was kept in an individual home cage, with water and grit freely available. With a few exceptions, the experiment was conducted once a day, seven days a week, at approximately the same time of day for each pigeon. The birds began each daily session at approximately 80% of their free-feeding weight. All birds had experience with matching-to-sample tasks.

Apparatus

Two LVE (Lehigh Valley Electronics) experimental chambers were used (two birds were run in each chamber). The chambers were 34-cm high, 35-cm long and 31-cm wide. On one of the walls of the chamber, three circular response keys, 2.5 cm in diameter, were arranged horizontally, 9 cm apart center-to-center. The bottom of each key was 22.5 cm above the wire mesh floor. Behind each key there was a 12-stimulus IEE (Industrial Electronics Engineers) in-line projector. In the same wall as the keys,

8.5 cm above the floor, a 6-cm wide, 5-cm high, horizontally-centered opening granted access to the food hopper. When the hopper was activated to provide mixed grain, a 28-V, 0.04-A lamp illuminated the opening. On the opposite wall, 30 cm above the floor, a 28-V, 0.1-A houselight provided general illumination to the chamber. An exhaust fan circulated air and masked outside noises.

In this experiment, the side keys were illuminated with red or green hues and the central key was illuminated with a white hue. The ABET II software (Lafayette Instrument Company) was used to control the experiment and record the data.

Procedure

Training. A trial began with the presentation of the sample stimulus (white hue on the center key) for 2, 6 or 18 s (henceforth referred to as S_2, S_6 and S_{18}, respectively). After the termination of the sample stimulus, two comparisons (red hue and green hue) were presented, one on each of the side keys. One comparison was correct following S_2 and S_{18}, and the other comparison was correct following S_6. The correct comparisons were counterbalanced across pigeons. A peck on one comparison turned off both keys, and reinforcement (access to mixed grain) was delivered if the choice was correct. To minimize feeding outside the experimental session, reinforcement duration was adjusted individually, and varied from 1 to 4 s across animals. Following reinforcement, a houselight-illuminated, 30-s intertrial interval (ITI) began. The houselight was illuminated only during the ITI. If the incorrect comparison was chosen, no reinforcement was delivered and the ITI started immediately.

To expedite learning, a correction procedure was used. If a response was incorrect, that trial was repeated. However, if three consecutive mistakes were made, in the following trial only the correct comparison would be presented. Excluding correction trials, a session comprised 64 trials, 16 x S_2 trials, 32 x S_6 trials, and 16 x S_{18} trials. This proportion of trial types ensured that both comparisons were correct the same number of times. Each comparison was presented the same number of times on each of the two side keys. Training continued until matching accuracy to each sample was at least of 80% for three consecutive sessions, or 60 sessions were reached.

Retention Test. Trials with a retention interval were added. The retention interval, spent in darkness, was introduced between sample offset and comparison onset, and could last 2.5, 5, 10 or 20 s. A session comprised 80 trials, 48 regular training trials (12 x 2 s, 24 x 6 s, 12 x 18 s) and 32 retention-interval test trials (16 following S_6, and 8 following S_2 and S_{18}). Each retention interval was presented the same number of times following each sample. Irrespective of trial type, correct responses were always reinforced, but the correction procedure was in effect only on regular trials. Reinforcement durations were maintained from Training. This phase ran for five sessions.

Retraining. Birds returned to the Training phase to ensure that baseline performance was maintained.

No-sample Test. Trials where no sample stimulus was presented were introduced. In these trials, the comparison stimuli were presented immediately following the end of the ITI. A session comprised 72 trials, 64 regular training trials (16 x 2 s, 32 x 6 s, 16 x 18 s) and 8 no-sample test trials. Responses on no-sample trials were never reinforced. Reinforcement durations were maintained from Training. This phase ran for five sessions.

Results

Training and Retraining. Only one of the four pigeons managed to reach the learning criterion before the maximum of 60 sessions was reached: P639 completed training in 21 sessions. The remaining three birds, after 60 sessions of training, were not consistently performing above the criterion of 80% correct responses: average matching accuracy on the last 5 sessions of training was 69% for S_2, 74% for S_6, and 81% for S_{18}. Since these values were significantly above chance (S_2: $t(2) = 9.82$, $p = .010$, $d_z = 5.67$, 95% CI for d [0.71, 10.90] ; S_6: $t(2) = 6.32$, $p = .024$, $d_z = 3.65$, 95% CI [0.29, 7.16] ; S_{18}: $t(2) = 6.42$, $p = .023$, $d_z = 3.70$, 95% CI [0.30, 7.26]), all birds proceeded to the Retention Test. The birds spent between 1 and 3 sessions ($\bar{x} = 2$) on the Retraining phase.

Retention Test. Figure 14 shows the results of the Retention Test. A repeated-measures ANOVA with sample duration (three levels) and retention interval

(five levels) as factors revealed no significant main effect of sample duration, $F(2, 6)$ = 2.73, $p = .143$, $\eta_G^2 = .26$, but a significant main effect of retention interval, $F(4, 12)$ = 6.71, $p = .004$, $\eta_G^2 = .19$. The interaction also was not significant, $F(8, 24) = 1.40$, $p = .248$, $\eta_G^2 = .14$, showing that the introduction of retention intervals affected matching performance, but there were no consistent differences among samples. Performance following 2-s samples appeared to be modulated by the retention interval: matching accuracy started by decreasing following the shorter retention intervals, and then increased, approaching the initial values following the longer retention intervals.
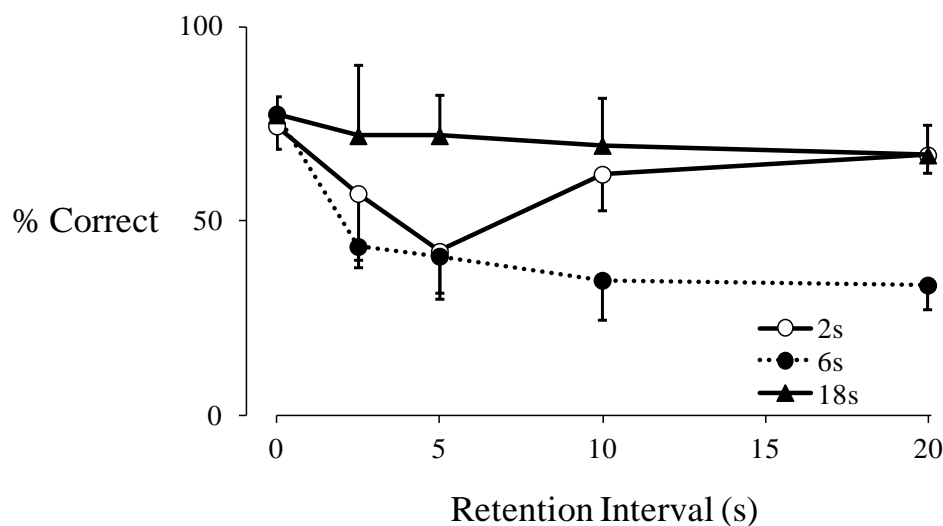
Figure 14. Mean (with SEM) percent correct following each of the three sample durations as a function of retention interval. 2-s and 18-s samples (filled lines) shared the correct comparison.

A difference between 6-s trials and 2- and 18-s trials seemed to emerge only for longer retention intervals. However, after the longest retention interval (20s), a paired-samples t-test showed only a marginally-significant effect: $t(3) = 3.00$, $p = .058$, $d_z = 1.5$, 95% CI: [-0.041, 2.964]. Only one pigeon, PG23, showed a clear difference between 6-s samples and the remaining samples in all retention intervals. For the remaining animals, matching accuracy did not differ much between samples,

but for the longest retention interval (20 s), accuracy for S_2 and S_{18} was consistently above S_6, a result that is reflected on the average function.

No-sample Test. All animals showed a strong preference for the comparison associated with 2- and 18-s samples (the "default" comparison): on average, that comparison was chosen 87.5% of the times. A t-test confirmed that group preference for the "default" option was significantly above 50%, $t(3) = 10.19$, $p = .002$, $d_z = 5.10$, 95% CI: [1.204, 9.095]. Individually, a two-tailed normal approximation to the binomial showed that preference for "default" was above chance ($\alpha = .05$) for all animals.

Discussion

In the present experiment, a task similar to studies I, II and III was used: All studies employed three samples and two comparisons, but while in this experiment the two extreme durations shared a comparison, in studies I, II and III the two longest samples shared a comparison. Figure 15 shows the sample-comparison mappings used in the studies so far. We were interested in assessing the effect of different sample-comparison mappings, by comparing the results with the previous studies in the dissertation. Additionally, since Singer et al. (2006) used the same sample-comparison mapping in a similar task, we were interested in replicating their findings.
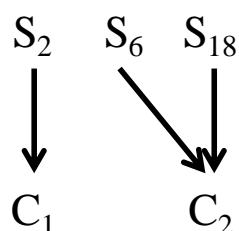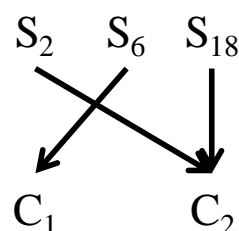

Figure 15. Mapping of samples (S) and comparisons (C) in the previous studies (left panel) and in the present study (right panel). The subscripts identify either the sample durations (2, 6 or 18) or the number of samples mapped to each comparison (1 or 2).

Regarding acquisition, after 60 sessions of training three birds did not meet the learning criterion. This result contrasts with studies I, II and III, in which pigeons completed training after an average of 23 sessions. Since the ratio between samples was maintained, the durations should be equally discriminable (Weber's law; see e.g., Gallistel, 1990; Gibbon, 1977). Therefore, if the animals were establishing three response rules, one per sample, differences in acquisition would not be expected. The finding that the sample-comparison mapping can make a task harder to acquire suggests that the animals were not treating samples equally in learning the task. If the animals were isolating the sample with an exclusive comparison, it would probably be easier to isolate one of the extreme durations instead of the intermediate duration. On the one hand, when isolating an extreme duration (S_2), two discriminations may be necessary: one easier (discriminating between S_2 and S_{18}) and one harder (discriminating between S_2 and S_6). On the other hand, when isolating the intermediate duration (S_6), the two discriminations (S_6 and S_2, S_6 and S_{18}) are difficult: both extremes are relatively close to the intermediate duration.

Alternatively, if the pigeons are learning the task by establishing thresholds between sample durations, the present experiment would also be more difficult to learn. In studies I, II and III, one comparison was correct following the shortest sample (2 s), and the other comparison was correct following the two longer samples (6 s and 18s). To learn that task, one threshold would be sufficient, between 2 and 6 s. On the other hand, in the present task, two thresholds would be needed: one between 2 and 6 s and another between 6 and 18 s. Since more thresholds would be necessary, this second task would be arguably harder to learn. In any case, the results are not consistent with the animals establishing three response rules, one per sample.

Singer et al. (2006) found a similar difference in acquisition: birds took an average of 56 sessions to learn their task in Experiment 2, but in Experiment 1, birds took an average of 33 sessions to learn a task where one comparison was correct following the shortest sample (0 s) and the other comparison was correct following the intermediate and longest samples (2 and 10 s). The difference in acquisition could be due to the different samples, but it could also be another example of how one mapping may be harder to learn than another.

In the Retention Test, Singer et al. (2006) found that, on the one hand, matching accuracy following the sample associated with an exclusive comparison decreased with retention interval. On the other hand, performance following the two samples that shared a comparison remained high at all retention intervals. Our results did not follow the same pattern: while the sample with an exclusive comparison (6 s) also decreased with retention interval, performance following the two samples that shared a comparison was not the same. Whereas matching accuracy for 18-s samples tended to stay at a high level, matching accuracy for 2-s samples fluctuated more, decreasing for the shorter retention intervals and then increasing for the longer retention intervals.

The pattern found is consistent with a hypothesis that has been put forward in previous studies of the dissertation, that animals are – at least in some of the trials – timing the interval since the end of the ITI. In that case, the timed interval would consist of the sum of sample and retention interval durations. For 2-s samples, that would mean that, for shorter retention intervals the total added duration would approach 6 s, which would lead to choices of the incorrect comparison. For longer retention intervals, the added duration would approach 18 s, which would lead to choices of the comparison correct following 18-s samples, which was also the correct comparison for 2-s samples, so matching accuracy should improve. For 6-s samples, the addition of the retention interval would lead to intervals greater than 6 s (and closer to 18 s), so as retention interval increased, incorrect responses (choices to the 18-s comparison) should also increase. For 18-s samples, the addition of the retention interval would only make the timed interval even longer, so the animal should continue to prefer the 18-s comparison. The effect can be easily seen in Figure 16, where data from each sample was replotted as a function of time since end of ITI. For the shortest durations, there is a preference for the 2- and 18-s comparison (C_2), as durations approach 6 s there is a decrease in this preference, which rises again when durations approach 18 s.

A difference that may be of importance between the present study and Singer et al. (2006) is in the salience of the stimuli. Singer et al. (2006) used two houselights, an unfiltered houselight as the sample, and a blue-filtered houselight to signal the ITI. We used the houselight to signal the ITI, and a white keylight as the sample. The

difference between stimulus saliences is arguably greater in our experiment, and that could have led the houselight, the most salient of the pair, to be used as a time marker.
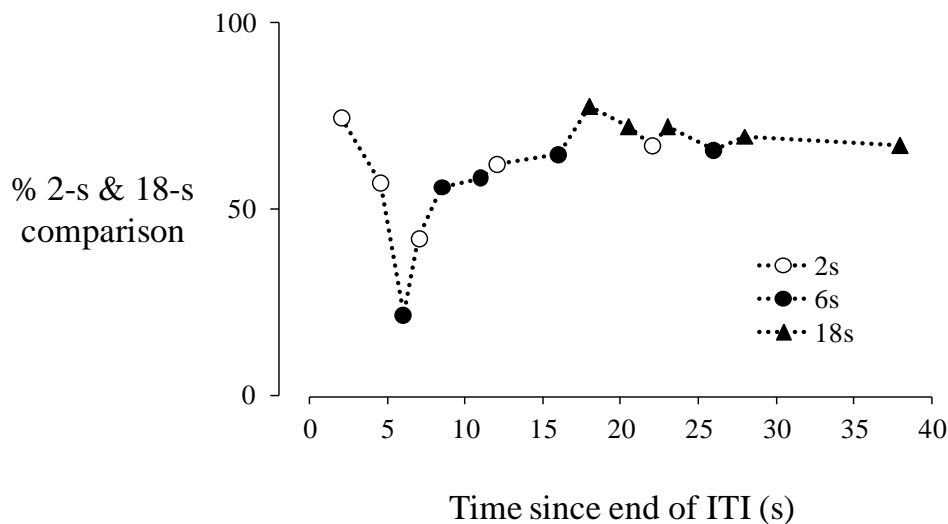
Figure 16. Mean percent of choices of the comparison associated with 2-s and 18-s samples in the Retention Test as a function of time since the end of the ITI. Data from all sample durations (signaled by different markers) are connected.

In sum, for the longer retention intervals, we found a preference for the shared comparison, a result consistent with Singer et al. (2006)'s findings. However, the cause behind this tendency may differ between the two studies. In Singer et al. (2006), the preference for the shared comparison could be either due to that comparison being the default choice (as per the single-code / default hypothesis), or due to a bias for the comparison associated with the shortest sample. In our experiment, the preference for the shared comparison could be due to the animals preferring the comparison associated with the longest sample (as a consequence of adding sample and retention interval).

Finally, the results of the no-sample test revealed a strong preference for the comparison shared between 2-s and 18-s samples, the "default" comparison. This preference is predicted by a single-code / default coding. However, this comparison is also the correct comparison following the shortest sample, which may be most

relevant to explain this result. In studies I and II, in retention testing animals showed a preference for the "default" comparison, which in that case was the comparison correct following 6- and 18-s samples. However, when a no-sample test was run, a clear, reliable preference for the comparison correct following 2-s samples was found. Therefore, in this no-sample test the animals may be showing a preference for a comparison because it is associated with the shortest sample duration, and not because it is the "default" choice, as the single-code / default hypothesis posits. Since the same no-sample test results were found in tasks with different mappings (and different preferences in retention testing), these results question the premise that the effect of a retention interval (losing information regarding the sample duration) is similar to having no sample presented.

In conclusion, the change in sample-comparison mapping in regards to studies I, II and III did not seem to affect what the animals learned. Even though the mapping used did not replicate Singer et al. (2006)'s findings, it provided additional support for some hypotheses advanced in other studies of the dissertation, namely that animals do not establish one response rule per sample when this type of task is learned, and that the houselight that signals the ITI can be a time marker.

CODING IN PIGEONS: BOUNDARY CONDITIONS OF THE SINGLE-CODE / DEFAULT STRATEGY IN A VISUAL DISCRIMINATION TASK

Eight pigeons learned a matching-to-sample task with three samples and two comparisons: one comparison was correct following two of the samples (many-to-one mapping), and the other comparison was correct following the remaining sample (one-to-one mapping). To solve the task, animals could establish three response rules, one per sample, or adopt a more economical coding strategy, the single-code / default strategy, which requires the establishment of only two response rules: one rule specific to the sample mapped one-to-one (the single code), and another rule to be applied following any other sample (the default rule). According to the single-code / default strategy, in retention testing animals should resort to the default rule and show a preference for the comparison that was matched many-to-one. However, with the exception of one pigeon, performance following the three samples was similar for most birds, a result that is not consistent with a single-code / default strategy. Our results suggest that the 3-sample version of task is insufficient for a single-code / default strategy to be broadly adopted.

By retrieving information from its environment, an animal can establish response rules (or codes) that will subsequently control its behavior. Flexibility in these response rules is instrumental in solving the challenges posed by an ever-changing environment. The more flexible, the more likely an animal is to come up with different ways to solve a problem. Among the possible ways to solve a specific problem, some solutions are likely to be more adequate (i.e., yielding better results) or more efficient (i.e., more economical in the resources it requires) than others.

The establishment of response rules (known as coding), as well as flexibility in selecting the best solutions for a problem have been topics of interest in the area of animal cognition, and a task commonly used to study these issues is the matching-to-sample task. In this task, following the presentation of a sample stimulus, there is a choice between two or more comparison stimuli. Choice of a specific comparison will be correct depending on the sample presented previously: for instance, in a task with samples S1 and S2 and comparisons C1 and C2, if S1 was presented, C1 should be selected, and if S2 was presented, C2 should be selected. In this case, learning the task could involve learning two sample-specific response rules: "If S1, choose C1" and "If S2, choose C2". Alternatively, the animals could have adopted a "single-code / default" strategy, establishing one sample-specific code (e.g., "If S1, choose C1"), and a non-specific, default code for all other samples (e.g., "If not S1, choose C2"). Following this strategy, C2 would always be chosen, unless information related with S1 was available at the moment of choice – C1 would be chosen in that case.

Based on the assumption that coding strategies depend on the particularities of each task, the adoption of a single-code / default strategy has been studied on a variety of matching tasks. For instance, when samples differ in salience, pigeons seem to establish a specific code only to the most salient sample (Grant, 2009b; Wixted & Gaitan, 2004). Similarly, when amount of training is varied between samples, it appears that a specific code is created only for the sample which was trained more extensively (Grant, 2006a; Grant & Blatz, 2004). Of most interest to the present work is a matching task that employs many-to-one mapping: one comparison (C1) is correct following two or more samples (e.g., S1, S2). By contrast, the other comparison (C2) is correct following only one sample (S3). In this case, instead of establishing three separate codes, one for each sample, a single-code / default strategy

based on two codes – "If S3, choose C2" (single code) and "If not S3, choose C1" (default) – may be more efficient.

The more samples used, the more advantageous the single-code / default strategy can be: In a task with three samples, if an animal adopts a single-code / default strategy (as opposed to sample-specific codes), it establishes two codes instead of three. In a task with, say, five samples (in which four share a correct comparison), the savings brought by a single-code / default strategy are more significant: two codes versus five. Therefore, the more samples used, the greater the incentive to form a common, "default" code.

There have been a few studies that have looked at whether animals would use a single-code / default strategy in the three-sample, two-comparison task. However, the evidence found remains unclear, either because of potential confounds related with the use of 0-s samples (Gaitan & Wixted, 2000; Zentall, Klein, & Singer, 2004) or because of other coding strategies also accounting for the results found (Singer, Klein, & Zentall, 2006). For a more in-depth discussion of these studies, please refer to the Introduction, p. 2. However, it is worth noting that the tasks employed in those studies have used the minimum number of samples (three), the situation in which the benefit of a single-code / default strategy is the smallest.

By contrast, the non-temporal discrimination that has provided the strongest evidence of the adoption of a single-code / default strategy has employed five samples (Clement & Zentall, 2000), a procedure more favourable to the adoption of this coding strategy. In the present study we explored the boundary conditions of the single-code / default strategy by assessing whether, in a non-temporal discrimination, a setup equivalent to the one used with temporal discriminations (3 samples, 2 comparisons) would be sufficient to occasion a single-code / default strategy.

Method

Subjects

Eight pigeons (*Columba livia*) were maintained at approximately 80% of their free-feeding body weight. The room where the animals were kept was maintained in a 13:11 hour light/dark cycle, with lights on at 08:00 and was kept between 20-22 ºC. In

each individual home cage water and grit were freely available. The experiment was conducted once a day, seven days a week, at approximately the same time of day for each pigeon. No bird had experience with non-temporal visual discrimination tasks.

Apparatus

Four experimental chambers were used (two birds were run in each chamber): Three LVE (Lehigh Valley Electronics) chambers and one homemade chamber. The LVE chambers were 34-cm high, 35-cm long and 31-cm wide. One of the walls of the chamber featured three circular response keys, 2.5 cm in diameter, arranged horizontally. The three keys were 9 cm apart, center to center, and the bottom of each key was 22.5 cm above the wire mesh floor. Behind each key, a 12-stimulus IEE (Industrial Electronics Engineers) in-line projector presented visual stimuli. Below the keys, a 6-cm wide, 5-cm high opening granted access to a food hopper with mixed grain. The hopper opening was centered horizontally, 8.5 cm above the floor, and was illuminated with a 28-V, 0.04-A light when the hopper was activated. On the opposite wall, 30 cm above the floor, a 28-V, 0.1-A houselight provided general illumination to the chamber. An exhaust fan circulated air and masked outside noises.

The homemade chamber had a similar setup to the LVE chambers. It measured 31 x 33 x 33 cm (h x l x w) and was equipped with three circular response keys, arranged horizontally. The keys, 2.5 cm in diameter and 21 cm above the wire mesh floor, were 9 cm apart (center to center). A 12-stimulus IEE (Industrial Electronics Engineers) in-line projector was installed behind each key. Access to a LVE food hopper was made through a 6-cm wide x 4.5-cm high horizontally-centered opening, located below the response keys, 6.5 cm above the floor. When the hopper was activated, a 28-V, 0.04-A light illuminated its opening. On the opposite wall, 27.5 cm above the floor, a 28-V, 0.1-A houselight provided general illumination. The operant chamber was enclosed by a PVC sound attenuating cubicle (Med Associates, ENV-018V) equipped with an exhaust fan.

In this experiment, the side keys were illuminated with vertical or horizontal white bars on a black background, and the central key was illuminated with red, green, blue or white hues. The ABET II software (Lafayette Instrument Company) was used to control the experimental events and record the data.

Procedure

Training. A trial started with the sample stimulus (a red, green, blue, or white hue) presented on the center key. Following five pecks on the center key, the sample was removed and the two side keys were illuminated with the comparison stimuli, a horizontal and a vertical bar. One response on a side key turned off both keys and, if the response was correct, was followed by access to food and then a 30-s intertrial interval (ITI), during which the houselight illuminated the box. An incorrect response was followed immediately by the ITI. When the ITI ended, the houselight was turned off and a new trial started. A correction procedure was used, in which a trial was repeated following an incorrect response. In the trial following three consecutive incorrect responses, only the correct comparison was presented. Each animal saw three different sample colors. One comparison was correct following two samples (S1 and S2), while the other comparison was correct following one sample (S3). The colors corresponding to each sample were counterbalanced across pigeons.

Excluding correction trials, a session was composed of 64 trials: 16 x S1 trials, 16 x S2 trials and 32 x S3 trials. This trial proportion assured that, in a session, each comparison was correct the same number of times. Each comparison was presented the same number of times on each of the lateral keys. Training continued for a minimum of 10 sessions, and until the animals were choosing the correct comparison following each sample at least 80% of the trials, for two consecutive sessions. To minimize feeding outside the session, reinforcement duration was adjusted individually, and varied from 2 s to 5 s across animals.

Retention Test. On some trials, a retention interval was introduced between sample offset and comparison onset (a period of darkness that could last 2.5, 5, 10 or 20 s). A session comprised 80 trials: 48 regular training trials (12 x S1, 12 x S2, 24 x S3) and 32 retention-interval test trials (8 x S1, 8 x S2, 16 x S3), randomly interspersed. Each retention-interval duration was presented the same number of times following each sample. In both training and retention-test trials, correct responses were always reinforced. The retention test lasted five sessions.

Results

The pigeons required from 10 to 25 sessions ($\bar{x} = 18$) to learn the task. Figure 17 shows the average results of the retention test. Performance following the three samples was similar, with matching accuracy falling to chance levels with the introduction of retention intervals.
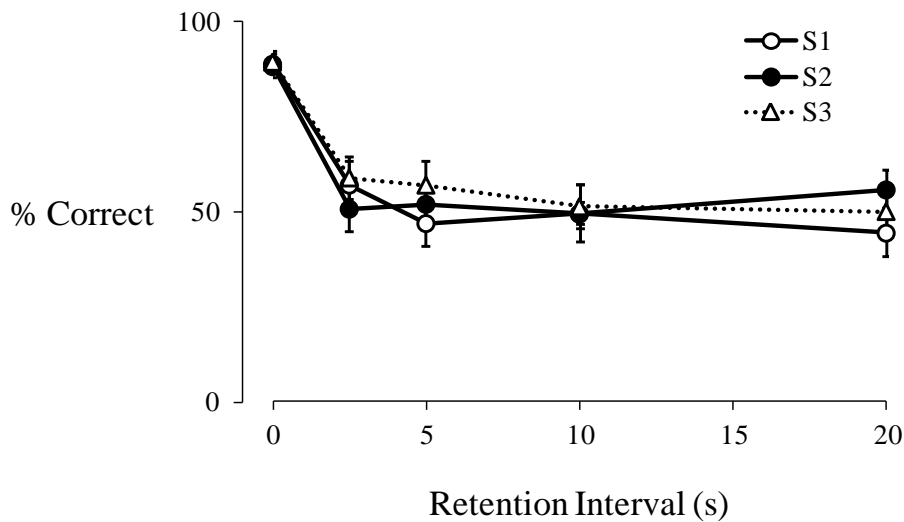

Figure 17. Mean (with SEM) percent correct to each of the three samples as a function of retention interval duration.

A repeated-measures ANOVA with sample (three levels) and retention interval duration (five levels) as factors revealed no significant main effect of sample, $F(2, 14) = .242$, $p = .788$, $\eta_G^2 = .014$ and a significant main effect of retention interval, $F(4, 28) = 56.44$, $p < .001$, $\eta_G^2 = .537$. The interaction was not significant, $F(8, 56) = .740$, $p = .656$, $\eta_G^2 = .039$. These results confirm that the retention interval had an effect on responding, but it did not differ between samples.

Additionally, the effect of the retention interval seemed to be independent of its duration; performance following all retention intervals was similar. This reading was confirmed by a repeated-measures ANOVA with sample (three levels) and non-zero retention interval duration (four levels) as factors, that revealed no significant main effect of either sample duration, $F(2, 14) = .253$, $p = .780$, $\eta_G^2 = .017$, or

retention interval, $F(3, 21) = 1.43$, $p = .262$, $\eta_G^2 = .021$. The interaction was also not significant, $F(6, 42) = .895$, $p = .508$, $\eta_G^2 = .038$. Taking the three samples together, percent correct following non-zero retention intervals was not significantly different from chance (95% Confidence Interval = 38% - 58%).

Figure 18 shows the individual performances on the retention test. The overall pattern is the same for most birds, as seen in the average function: the introduction of retention intervals leads to a decrease in matching accuracy, and there does not seem to be a clear preference for any comparison. There is, however, one exception: PG45 (bottom left panel), showed a preference for the comparison mapped with S1 and S2. For this pigeon, performance following S1 and S2 stayed at a high level, while performance following S3 decreased markedly. To illustrate, following the longest retention interval (20 s), percent correct to S1 and S2 was 70%, whereas percent correct to S3 was only 20%. The preference for the comparison mapped many-to-one is expected if a single-code / default strategy had been employed.
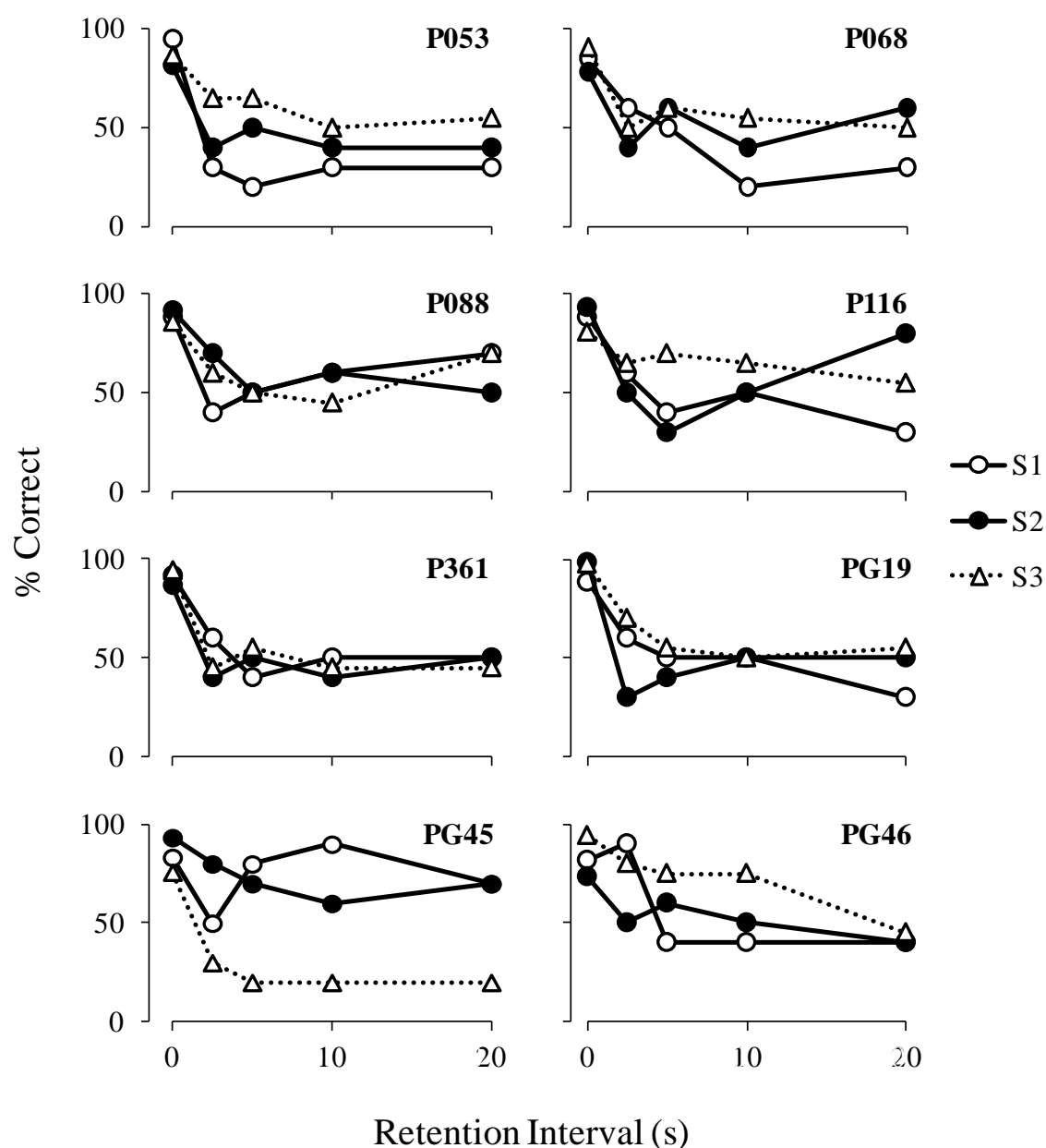
Figure 18. Percent correct to each of the three samples as a function of retention interval for each bird.

Discussion

In this study, we employed a matching-to-sample task with three samples and two comparisons; one comparison was correct following two samples (many-to-one mapping) and the other was correct following the remaining sample (one-to-one-

mapping). This task has been commonly used to study the adoption of a single-code / default strategy in temporal discrimination tasks. The strategy consists in establishing two response rules, one specific to the sample mapped one-to-one, and a general, default rule applied to any other sample. A single-code / default strategy may be advantageous because the establishment of a response rule that applies to more than one sample should be more economical than establishing one response rule per sample.

The more samples share a correct comparison, the greater the savings brought by the adoption of a general response rule. Therefore, the 3-sample task is a boundary condition, the case in which the benefits of this strategy are smallest. When only two samples share a comparison, the adoption of a general rule reduces only one rule (going from two sample-specific rules to one general rule that would apply to both samples). Some results obtained with this task have been suggestive that a single-code / default strategy may be adopted in timing tasks (Singer, Klein, & Zentall, 2006; Study I), but the data has not been conclusive.

In non-temporal discriminations, Clement and Zentall (2000) found evidence suggesting that a single-code / default strategy can be used, but more than two samples were matched many-to-one, so the task offered greater benefits to the adoption of this coding strategy. The purpose of the present experiment was to assess whether the 3-sample version of the task would be sufficient to lead to the adoption of a single-code / default strategy in a non-temporal discrimination.

The average results of the Retention Test were not consistent with the single-code / default strategy: all samples were similarly affected by the retention intervals. That is, a "default" response rule triggered by the retention interval seemingly was not established. Therefore, it seems that, overall, the 3-sample task was insufficient for the arguably more efficient single-code / default strategy to be adopted. There was, however, one exception. Pigeon PG45 showed the preference pattern predicted if a single-code / default strategy was used: Following a retention interval, the animal was likely to resort to the "default" response, that is, to choose the comparison associated with S1 and S2. Hence, in retention testing, performance following S1 and S2 remained at a high level while performance following S3 declined.

If PG45 was the only animal that resorted to the more economical coding strategy, it would be expected that acquisition would be fastest for this animal. In fact, PG45 was the fastest pigeon to learn the task, taking only 10 sessions to meet criterion. To illustrate, the second-fastest bird took 16 sessions to complete training.

Even when faced with the same task, different animals may sometimes attend to different stimuli or learn different things (e.g., Reynolds, 1961; Gaitan & Wixted, 2000; see also studies III, IV and V). Taking into account that different animals may approach tasks differently or be sensitive to different characteristics of a task, perhaps a better research question would ask how common a coding strategy would be in a given task. That is, when looking at how animals solve a task, instead of a single common solution, it is perhaps more adequate to adopt a plural view, with different solutions, some more likely to be adopted than others.

Animals are flexible in how they learn a task; for instance, depending on task difficulty, animals may adopt a prospective or a retrospective coding strategy (Zentall, Urcuioli, Jagielo, & Jackson-Smith, 1989). The present task may be another example of coding flexibility: even though the task allowed for a coding strategy that would be more economical and easier to learn, only one bird out of eight made use of it. By contrast, in a version of this task with five samples (four samples mapped to the same comparison), there was strong evidence suggesting that a single-code / default strategy was adopted (Clement & Zentall, 2000). The more samples share a comparison, the greater the savings a common code brings, so as the number of samples matched to the same comparison increases, the adoption of a common code may become more likely.

Therefore, a task employing four samples (with three samples sharing a correct comparison) should, in comparison with the 3-sample task, result in a higher proportion of animals adopting a single-code / default strategy. Also of relevance to this topic would be knowing if, following training in the 3-sample task used in this study, the introduction of the fourth sample would be sufficient for the pigeons to solve the task using a default rule. Would it be the same as training with four samples from the start? Perhaps the introduction of a new sample would simply lead to the creation of a new sample-specific rule, instead of changing the overall coding strategy. The more data collected on the conditions, boundaries, and dynamics of

adoption of coding strategies, the better will our understanding be of how animals adapt to and solve different tasks.

Discrimination learning can be conceptualized as involving two steps, first identifying the relevant stimuli and stimuli dimensions, and then establishing responses to those stimuli (Sutherland & Mackintosh, 1971). In the present dissertation, we studied both of these processes in matching-to-sample tasks, to identify the stimuli attended to and, based on those stimuli, the kind of response rules developed.

On the sources of stimulus control

Regarding the stimuli used attended to, in matching-to-sample tasks, the sample is the one usually assumed to exert the greatest influence on choice. The results of Study I were consistent with the intertrial (ITI) illumination in our task being a source of stimulus control. Study II supported the hypothesis that the ITI houselight offset was a time marker, jointly with the sample. Study III suggested that there was a trade-off in the usage of the two sources of stimulus control, which differed among animals. The retention functions in Study IV were also consistent with responding based – at least partially – on the time elapsed since the ITI. Therefore, in our temporal discrimination tasks, timing appeared to be initiated by the ITI houselight offset and/or the sample keylight onset.

Having two sources of stimulus control brings up some interesting questions regarding how the relative importance of each stimulus is established and how it may change with experience. For instance, does control by different stimuli emerge together, or do some stimuli acquire control over responding quicker than others? What determines which stimulus develops the most control over responding? Does each animal have its own particular predisposition or is it a random process and, whatever happens to be the stimulus the animal attends to and delivers the most reinforcement in early stages of learning the task becomes predominant? Running tests to assess the relative importance of each stimulus (as in Study III) in different phases of acquisition may help address some of these questions.

The contents of temporal discrimination learning

As discussed above, our results suggested that two different stimuli controlled responding. Taking into account that the retention tests were designed to assess only the coding strategies related to the nominal sample (keylight), whenever the animals were responding based on stimuli other than the sample, the analysis of results must proceed with some caution. More specifically, we assume that when timing was initiated by the houselight offset, the animals were timing an empty interval, and the stimulus that terminated the to-be-timed interval was the presentation of the comparisons. The introduction of retention intervals would increase the interval between houselight offset and comparison offset, so the animals would perceive retention-interval trials as trials with longer sample durations. In that case, a retention test would not be informative of the coding strategies underlying choice, because, irrespective of coding strategy, the introduction of retention intervals should always lead to choices of the "long" key.

Therefore, retention-test performance should be a combination of two effects: a preference for the comparison associated with the longest sample in houselight-controlled trials, and a reflection of the coding strategy in sample-controlled trials. In the case of single-code / default, there should be a preference for the comparison mapped many-to-one, which in studies I, II and IV happened to include the longest sample. So, the two effects would predict similar results: preference for the key associated with 6- and 18-s samples in studies I and II, and preference for the key associated with 2- and 18-s samples in study IV.

In the case of multiple coding, in retention testing there should be a preference for the comparison associated with the shortest sample, a result at odds with the predictions of single-code / default in studies I and II, but not in Study IV. Hence, in studies I and II, if the animals adopted multiple coding, they would choose the "short" comparison when responding was under sample control, and the "long" comparison when responding was under houselight control. The overall preference would depend on the degree of control of each stimulus over responding, but assuming that both stimuli equiprobably control behavior, on average they should cancel each other out; a strong preference for the "long" comparison would be unlikely in group data. Thus, the tendency for choosing the "long" key found in studies I and II would be more likely with single-code / default.

In contrast, the results of Study V do not support the hypothesis that animals were resorting to single-code / default. In that study, a 3-sample, 2-comparison non-

temporal discrimination task was insufficient for a widespread adoption of that strategy. However, it may be the case that the necessary conditions for the adoption of a coding strategy may differ based on the modality of the stimuli.

It is worth noting that the mapping used in studies I and II could have induced an alternative coding strategy, consisting in the establishment of a threshold between 2 s and 6 s, and responding being based on whether the sample duration was below or above the threshold. Even though this would not be a case of single-code / default it would be a different example of coding flexibility, where three sample-specific codes would be replaced by two codes ("If below threshold, choose C1" and "If above threshold, choose C2"). Both coding strategies would involve grouping of samples: in single-code / default it would be grouping all the samples that are not mapped to an exclusive comparison; in the threshold strategy it would be grouping all samples above a certain value.

The finding that, when sample-comparison mapping was changed in Study IV, learning was much slower than in the previous studies, suggests that the three sample durations were not treated equally during learning, which could reveal that the pigeons were grouping or isolating some of the samples. Overall, even if our data are not able to provide a definite identification of the coding strategies in use, they do suggest that animals did not create a response rule per sample duration, which is a demonstration of flexibility in solving these tasks.

Is a delay always a retention interval?

The insertion of a delay before a response has a long tradition as a means to study memory; one of the first mentions of this practice is over a century old (Hunter, 1913). However, a delay may not always be an adequate means to assess memory. In a delayed matching-to-sample task, the delay between sample and comparisons is a period where the subject must maintain information necessary to select the correct comparison – hence being also known as a retention interval. To that end, during the delay the stimulus dimension that was trained is not presented. For instance, in a visual discrimination, the delay may be a period of darkness so that no other visual stimuli interfere with the performance.

However, in temporal discrimination tasks, such a control is not attainable; it is impossible to remove the relevant dimension (time) from the delay. In fact, a delay may be interpreted as a new temporal stimulus instead of a retention interval. In a delayed

matching-to-sample task, a retention interval is assumed to eventually lead to forgetting, which would put the animal in a situation akin to having no sample presented in the first place. The difference in preferences we found in retention and no-sample trials may suggest that these assumptions may not be true; the delay may not be testing memory, or the effect of forgetting may not be the same as having no sample presented in the first place. Overall, in all our no-sample tests there was a preference for the comparison associated with the shortest sample, irrespective of whether a similar preference was shown in retention testing. That is, performance in the two tasks appeared to be unrelated.

It would also be of interest to know why, on no-sample trials, pigeons showed a preference for the "short" key. One possibility is that, while learning the task, during the ITI the animals may have been engaged in an activity that precluded them of attending to the intelligence panel quickly enough to see the sample. The sample that has the shortest duration is the most likely to be missed by the animal. Hence, the animals may have learned that when they do not see a sample, the "short" comparison is usually correct.

Closing remarks

In the present dissertation we aimed to identify the codes pigeons established to learn a 3-sample, 2-comparison matching-to-sample task. In that process, we came across several results of interest. Whereas in a 2-sample, 2-comparison task pigeons seem to establish a code per sample (Grant & Spetch, 1994), we found that pigeons did not treat the three samples equally, a sign that their coding strategies were adapted to the particularities of the task. We also found that more than one stimulus could control responding, and that the relative control of each stimulus could be quantified. Besides showing flexibility in how a task is learned, different animals may solve the same task differently – in Study III we saw significant individual differences in our estimations of reliance on different stimuli, and in several studies some animals appeared to have learned something different from the rest of the animals in the same task.

The present work illustrates the difficulty in unequivocally identifying the response rules controlling behavior, either because it is challenging to isolate the predictions of different hypotheses, or because small changes in procedure may lead to big changes on what is learned. Take as an example the ways intertrial and retention intervals can be signaled in a task: when they are similarly signaled, a risk of potential

confusion is introduced (e.g., Sherburne, Zentall, & Kaiser, 1998). When they are signaled differently, if the most salient stimulus is used during the ITI, the ITI itself may become a discriminative stimulus the animals use to learn a task, as we saw in the present work. If the retention interval is signaled with a salient stimulus, such as a houselight, it may cause interference and disrupt performance (e.g., Grant & Roberts, 1976; Harper & White, 1997; Roberts & Grant, 1978).

These observations call attention to the potential disconnect between the intentions of the experimenter when designing a task and what an animal learns from that task. For instance, as discussed previously, the nominal sample may not always be the effective sample; the effect of a retention interval may not coincide with the experimenter expectations, so care should be taken in the interpretation of what animals learn from the tasks they are faced with. As Murray Sidman recently put it:

> "All stimuli are, of course, inherently complex. If our subject responds to a key that has a line on it, we cannot tell on any particular occasion whether the response is controlled by the line's tilt, height, width, illuminance, distance from the edge of the key, or any of many other possibilities. Any single instance of stimulus control is, therefore, always an inference; the identification of a stimulus control topography requires many observations and many variations of possibly relevant stimulus features. We have to beware, especially, of attributing our own stimulus control topographies to our subjects and students." (Sidman, 2008, p.133).

The assumption that learning may be based on several stimuli, and that different animals may learn the same task in different ways may make the interpretation of behavior more challenging, but may also provide a step forward in our understanding of animal learning and behavior.

Blough, D. S. (1969). Attention shifts in a maintained discrimination. *Science, 166*, 125–126.

Carvalho, M. P., & Machado, A. (2012). Relative versus absolute stimulus control in the temporal bisection task. *Journal of the Experimental Analysis of Behavior, 98*, 23–44.

Catania, A. C. (1970). Reinforcement schedules and the psychophysical judgments: a study of some temporal properties of behavior. In W. N. Schoenfeld (Ed.), *The theory of reinforcement schedules* (pp. 1–42). New York: Appleton-Century-Crofts.

Cheng, K., Spetch, M. L., & Miceli, P. (1996). Averaging temporal duration and spatial position. *Journal of Experimental Psychology: Animal Behavior Processes, 22*, 175–182.

Church, R. M., & Deluty, M. Z. (1977). Bisection of temporal intervals. *Journal of Experimental Psychology: Animal Behavior Processes, 3*, 216–228.

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

Colwill, R. M. (1984). Disruption of short-term memory for reinforcement by ambient illumination. *The Quarterly Journal of Experimental, 36B*, 235–258.

Dorrance, B. R., Kaiser, D. H., & Zentall, T. R. (2000). Event-duration discrimination by pigeons: The choose-short effect may result from retention-test novelty. *Animal Learning & Behavior, 28*, 344–353.

Gaitan, S. C., & Wixted, J. T. (2000). The role of "nothing" in memory for event duration in pigeons. *Animal Learning & Behavior, 28*, 147–161.

Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: Bradford Books/MIT Press.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review, 84*, 279–325.

Grant, D. S. (1991). Symmetrical and asymmetrical coding of food and no-food samples in delayed matching in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 17*, 186–193.

Grant, D. S. (2001). Memory for empty time intervals in pigeons. *Animal Learning & Behavior*, *29*, 293–301.

Grant, D. S. (2006a). Asymmetrical sample training and asymmetrical retention functions in one-to-one and many-to-one matching in pigeons. *Learning and Motivation, 37*, 209–229.

Grant, D. S. (2006b). Training delays reduce the choose-short effect with keylight duration samples in pigeons. *Behavioural Processes, 72*, 139–148.

Grant, D. S. (2009a). Choose-long errors at delays shorter than the training delay persist when enhanced between-trial summation of duration memories are precluded in pigeons. *Behavioural Processes, 82*, 335–339.

Grant, D. S. (2009b). Use of samples differing markedly in salience may encourage use of a single-code / default strategy in matching-to-sample in pigeons. *Learning and Motivation, 40*, 386–393.

Grant, D. S., & Blatz, C. (2004). Asymmetrical sample training produces asymmetrical retention functions in feature-present/feature-absent matching in pigeons. *Learning and Motivation, 35*, 53–70.

Grant, D. S., & Roberts, W. A. (1976). Sources of retroactive inhibition in pigeon short-term memory. *Journal of Experimental Psychology: Animal Behavior Processes*, *2*, 1–16.

Grant, D. S., & Spetch, M. L. (1993). Analogical and nonanalogical coding of samples differing in duration in a choice-matching task in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 19*, 15–25.

Grant, D. S., & Spetch, M. L. (1994). The role of asymmetrical coding of duration samples in producing the choose-short effect in pigeons. *Learning and Motivation*, *25*, 413–430.

Grant, D. S., Spetch, M. L., & Kelly, R. (1997). Pigeons' coding of event duration in delayed matching-to-sample. In C.M. Bradshaw & E. Szabadi (Eds.), *Advances in Psychology (Vol 120): Time and Behavior: Psychological and Neuro-behavioral Analyses* (pp. 217–264). Amsterdam: Elsevier/ North Holland.

Grant, D. S., & Talarico, D. C. (2004). Processing of empty and filled time intervals in pigeons. *Learning & Behavior*, *32*, 477–90.

Harper, D. N., & White, K. G. (1997). Retroactive interference and rate of forgetting in delayed matching-to-sample performance. *Animal Learning & Behavior*, *25*, 158–164.

Hunter, W. S. (1913). The delayed reaction in animals and children. *Animal Behavior Monographs, 2*, 1–86.

Johnson, D., & Cumming, W. (1968). Some determiners of attention. *Journal of the Experimental Analysis of Behavior, 11*, 157–166.

Kelly, R., & Spetch, M. L. (2000). Choice biases in delayed matching-to-sample duration with pigeons: Manipulations of ITI and delay illumination. *The Quarterly Journal of Experimental Psychology, 53B*, 309–323.

Kirkpatrick, K., & Church, R. M. (2000). Stimulus and temporal cues in classical conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, *26*, 206–19.

Kirkpatrick-Steger, K., & Wasserman, E.A. (1996). The what and where of the pigeon's processing of complex visual stimuli. *Journal of Experimental Psychology: Animal Behavior Processes, 22*, 60–67.

Kraemer, P. J., Mazmanian, D. S., & Roberts, W. A. (1985). The choose-short effect in pigeon memory for stimulus duration: Subjective shortening versus coding models. *Animal Learning & Behavior, 13*, 349–354.

Kraemer, P. J., Mazmanian, D. S., & Roberts, W. A. (1987). Simultaneous processing of visual and spatial stimuli in pigeons. *Animal Learning & Behavior, 15*, 417–422.

Kraemer, P. J., & Roberts, W. A. (1985). Short-term memory for simultaneously presented visual and auditory signals in the pigeon. *Journal of Experimental Psychology: Animal Behavior Processes, 11*, 137–152.

Leak, T. M., & Gibbon, J. (1995). Simultaneous timing of multiple intervals: Implications of the scalar property. *Journal of Experimental Psychology: Animal Behavior Processes*, *21*, 3–19.

Leith, C. R., & Maki, W. S. (1975). Attention shifts during matching- to-sample performance in pigeons. *Animal Learning & Behavior, 3*, 85–89.

Lovejoy, E. (1968). *Attention in discrimination learning*. San Francisco: Holden-Day.

Mackintosh, N.J. (1965). Selective attention in animal discrimination learning. *Psychological Bulletin, 64*, 124–150.

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review, 82*, 276–298.

Maki, W. S., & Leith, C. R. (1973). Shared attention in pigeons. *Journal of the Experimental Analysis of Behavior, 16*, 345–349.

McIlvane, W., & Dube, W. (2003). Stimulus control topography coherence theory: Foundations and extensions. *The Behavior Analyst, 26*, 195–213.

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes, 9*, 320–334.

Meck, W. H., & Church, R. M. (1984). Simultaneous temporal processing. *Journal of Experimental Psychology: Animal Behavior Processes*, *10*, 1–29.

Pinto, C., Fortes, I., Jozefowiez, J., & Machado, A. (2012). Interaction of temporal memories in extinction. In P. Guilhardi, M. Menez, & F. López (Eds.), *Tendencias en el estudio contemporáneo de la estimación temporal* (pp. 193–222). Ciudad de México: Universidad Nacional Autónoma de México.

Pinto, C., & Machado, A. (2011). Short-term memory for temporal intervals: Contrasting explanations of the choose-short effect in pigeons. *Learning and Motivation, 42*, 13–25.

Reynolds, G. (1961). Attention in the pigeon. *Journal of the Experimental Analysis of Behavior, 4*, 203–208.

Roberts, W. A., & Grant, D. S. (1978). An analysis of light-induced retroactive inhibition in pigeon short-term memory. *Journal of Experimental Psychology: Animal Behavior Processes*, *4*, 219–236.

Roberts, W. A., & Mitchell, S. (1994). Can a pigeon simultaneously process temporal and numerical information? *Journal of Experimental Psychology: Animal Behavior Processes, 20*, 66–78.

Santi, A., Bridson, S., & Ducharme, M. (1993). Memory codes for temporal and nontemporal samples in many-to-one matching by pigeons. *Animal Learning & Behavior*, *21*, 120–130.

Santi, A., Hornyak, S., & Miki, A. (2003). Pigeons' memory for empty and filled time intervals signaled by light. *Learning and Motivation*, *34*, 282–302.

Santi, A., Ross, L., Coppa, R., & Coyle, J. (1999). Pigeons' memory for empty time intervals marked by visual or auditory stimuli. *Animal Learning & Behavior*, *27*, 190–205.

Sherburne, L. M., & Zentall, T. R. (1993). Coding of feature and no-feature events by pigeons performing a delayed conditional discrimination. *Animal Learning & Behavior, 21*, 92–100.

Sherburne, L. M., Zentall, T. R., & Kaiser, D. H. (1998). Timing in pigeons: The choose-short effect may result from pigeons' "confusion" between delay and intertrial intervals. *Psychonomic Bulletin & Review, 5*, 516–522.

Sidman, M. (2008). Reflections on stimulus control. *The Behavior Analyst, 31*, 127–135.

Singer, R. A., Klein, E. D., & Zentall, T. R. (2006). Use of a single-code / default strategy by pigeons to acquire duration sample discriminations. *Learning & Behavior, 34*, 340–347.

Skinner, B. F. (1984). The evolution of behavior. *Journal of the Experimental Analysis of Behavior, 41*, 217–222.

Spetch, M. L. (1987). Systematic errors in pigeons' memory for event duration: Interaction between training and test delay. *Animal Learning & Behavior*, *15*, 1–5.

Spetch, M. L., & Rusak, B. (1989). Pigeons' memory for event duration: Intertrial interval and delay effects. *Animal Learning & Behavior. 17*, 147–156.

Spetch, M. L., & Rusak, B. (1992). Temporal context effects in pigeons' memory for event duration. *Learning and Motivation, 23*, 117–144.

Spetch, M. L., & Wilkie, D. M. (1982). A systematic bias in pigeons' memory for food and light durations. *Behavior Analysis Letters, 2*, 267–274.

Spetch, M. L., & Wilkie, D. M. (1983). Subjective shortening: A model of pigeons' memory for event duration. *Journal of Experimental Psychology: Animal Behavior Processes, 9*, 14–30.

Stubbs, A. (1968). The discrimination of stimulus duration by pigeons. *Journal of the Experimental Analysis of Behavior, 11,* 223–238.

Stubbs, A. (1976). Scaling of stimulus duration by pigeons. *Journal of the Experimental Analysis of Behavior, 26*, 15–25.

Sutherland, N. S., & Mackintosh, N. J. (1971). *Mechanisms of animal discrimination learning*. San Diego, CA: Academic Press.

Sutton, J. E., & Roberts, W. A. (1998). Do pigeons show incidental timing? Some experiments and a suggested hierarchical framework for the study of attention in animal cognition. *Behavioral Processes, 44,* 263–275.

Thomas, D. R. (1970). Stimulus selection, attention, and related matters. In J. H. Reynierse (Ed.), *Current issues in animal learning* (pp. 311–356). Lincoln, NE: University of Nebraska Press.

Urcuioli, P. J., Zentall, T. R., Jackson-Smith, P., & Steirn, J. N. (1989). Evidence for common coding in many-to-one matching: Retention, intertrial interference, and transfer. *Journal of Experimental Psychology: Animal Behavior Processes, 15*, 264–273.

Vyazovska, O. V., Teng, Y., & Wasserman, E. A. (2014). Attentional tradeoffs in the pigeon. *Journal of the Experimental Analysis of Behavior, 101*, 337–354.

Weaver, J. E., Dorrance, B. R., & Zentall, T. R. (1999). Present/absent sample matching in pigeons: Is comparison choice controlled by the sample stimulus or by differential sample responding? *Animal Learning & Behavior, 27*, 288–294.

Weiss S.J. (1964). Summation of response strengths instrumentally conditioned along a composite-stimulus continuum during free-operant summation. *Journal of the Experimental Analysis of Behavior, 68*, 151–155.

Wilkie, D. M. (1978). Delayed symbolic matching to sample in the pigeon. *The Psychological Record, 28*, 463–469.

Wilson, B., & Boakes, R. A. (1985). A comparison of the short-term memory performances of pigeons and jackdaws. *Animal Learning & Behavior, 13*, 285–290.

Wixted, J., & Gaitan, S. (2004). Stimulus salience and asymmetric forgetting in the pigeon. *Learning & Behavior, 32*, 173–182.

Wolf, M. M. (1963). Some effects of combined S^Ds. *Journal of the Experimental Analysis of Behavior, 6,* 343–347.

Zeaman, D., & House, B. J. (1963). The role of attention in retardate discrimination learning. In N. R. Ellis (Ed.), *Handbook of mental deficiency: Psychological theory and research* (pp. 159–223). New York: McGraw-Hill.

Zentall, T. R., Klein, E., & Singer, R. (2004). Evidence for detection of one duration sample and default responding to other duration samples by pigeons may result from an artifact of retention-test ambiguity. *Journal of Experimental Psychology: Animal Behavior Processes, 30*, 129–34.

Zentall, T. R., Sherburne, L. M., Steirn, J. N., Randall, C. K., Roper, K. L., & Urcuioli, P. J. (1992). Common coding in pigeons: Partial versus total reversals of one-to-many conditional discriminations. *Animal Learning & Behavior*, *20*, 373–381.

Zentall, T. R., Steirn, J. N., Sherburne, L. M., & Urcuioli, P. J. (1991). Common coding in pigeons assessed through partial versus total reversals of many-to-one

conditional and simple discriminations. *Journal of Experimental Psychology: Animal Behavior Processes, 17*, 194–201.

Zentall, T. R., Urcuioli, P. J., Jagielo, J. A., & Jackson-Smith, P. (1989). Interaction of sample dimension and sample-comparison mapping on pigeons' performance of delayed conditional discriminations. *Animal Learning & Behavior, 17*, 172–178.