

Pigeons' performance in a tracking change-signal procedure is consistent with the independent
horse-race model

Stephen E. G. Lea, Pizza Ka Yee Chow, Christina Meier, & Ian P. L. McLaren

University of Exeter

& Frederick Verbruggen

University of Exeter and Ghent University

Manuscript XAN-2018-1496R2, under revision for the Journal of Experimental Psychology:
Animal Learning and Cognition. Please do not cite or quote without permission from an
author.

© Copyright 2019 S. E. G. Lea, P. K. Y. Chow, C. Meier, I. P. L. McLaren & F. Verbruggen

Correspondence concerning this article should be addressed to S. E. G. Lea, University of
Exeter, Department of Psychology, Washington Singer Laboratories, Exeter EX4 4QG, United
Kingdom. E-mail: s.e.g.lea@exeter.ac.uk

Author Notes

Stephen E. G. Lea, Pizza Ka Yee Chow, Christina Meier, and Ian P. L. McLaren,,
Department of Psychology, University of Exeter, UK

F. Verbruggen, Department of Psychology, University of Exeter, UK, and Department
of Experimental Psychology, Ghent University, Belgium.

Pizza Chow is now at the School of Environmental Science, Hokkaido University,
Japan; Christina Meier is now at the Department of Psychology, University of Manitoba,
Canada

The authors declare that they have no conflict of interest.

Data from the experiment described in this paper were presented at the meeting of
the Comparative Cognition Society held in Melbourne, Florida in April 2018. We are grateful
for comments and suggestions made by colleagues present at that meeting.

Abstract

In many cognitive tasks where humans are thought to rely on executive functioning, pigeons' behavior can be explained by associative processes. A key form of executive functioning is inhibiting prepotent responses, often investigated in humans by means of "Stop-signal" or "Change-signal" procedures. In these procedures, execution of a well-practised ("Go") response to a stimulus is occasionally interrupted by a signal to withhold or alter the practised response. Performance in such tasks is usually described by the "independent horse race model" model. This model assumes that the processes that cause the Go and inhibitory responses occur independently; the process that finishes first determines the response observed. We further tested this model by training pigeons to track the circular movement of a colored patch around a touchscreen by pecking it; the spot occasionally deviated from its normal path (the Change signal). The pigeons had to inhibit the habitual movement of their heads in order to land a peck on the spot in its unexpected position. The key predictions of the independent horse-race model were confirmed in the pigeons' latency data. Thus, the independent race model can also successfully describe Stop-change performance of subjects that do not rely on executive control.

Keywords: Executive control; stop-signal procedure; response latency; independent horse-race model; pigeons

Introduction

The processes underlying the inhibition of prepotent responses have been the subject of much recent research across disciplines (Verbruggen et al., 2019). In human cognition, response inhibition is seen as a core feature of executive control (Shallice & Burgess, 1993; Verbruggen & Logan, 2008b), and its taxonomic distribution has therefore been studied extensively (MacLean et al., 2014), with the object of understanding whether executive control can be attributed to any non-human species, and if so, which. However, response inhibition need not involve executive processes; indeed, inhibition has been a core element in theories of associative learning from Pavlov (1927) onwards. Evidence is beginning to accumulate suggesting that tasks commonly used to assess response inhibition can be understood in associative terms. For example, Meier, Lea and McLaren (2018) have recently demonstrated that pigeons can be trained in tasks modelled on the Stop Signal and Change Signal tasks commonly used to examine response inhibition in human cognition, and that their behavior in these tasks can be predicted by the independent horse-race model of Logan & Cowan (1984; see also Verbruggen & Logan, 2008b, 2009). The present paper is a further exploration of pigeons' performance in Change-Signal tasks.

In the Stop or Change-Signal task, the participant repeatedly has to make a well-practised response to a signal (the Go Stimulus), but occasionally this signal is either replaced, or almost immediately followed, by a different signal which requires the participant to withhold the normal response (in the case of a Stop Signal) or switch to a different one (in the case of a Change Signal). These test trials are referred to as Stop or Change trials. For example, in a Stop-Signal task, human participants might be instructed that a green dot will repeatedly be presented on screen, and that they should mouse-click on this green dot as soon as it appears. However, on some trials, the green dot might change to a red color shortly after the start of the trial. If this happens, the participants should withhold any response and instead wait for the trial to end. In a Change-Signal task,

participants might be instructed, when seeing the red dot, not to mouse-click on that dot but instead to move the mouse and click on a different target.

Logan and Cowan's (1984) independent horse-race model is a special case of a very general class of models of choice reaction times, in which different stimuli are thought to initiate independent mental or neural processes, with the outcome of the choice depending on which of those processes terminates first (see also Logan, Van Zandt, Verbruggen, & Wagenmakers, 2014). Other examples include the stochastic model of Audley (1960) and the Sequential Choice Model of Shapiro, Siller and Kacelnik (2008) – indeed, the latter authors also use the horse race metaphor. These models have been applied to a wide range of experimental situations, for example mid-session reversal in pigeons (Smith, Zentall & Kacelnik, 2018). The horse-race model specifically postulates that initiating a response and withholding a response are two independent processes: an excitatory go process that is triggered by the presentation of a stimulus demanding the response, and an inhibitory stop process triggered by a signal not to respond. Performance depends on the outcome of the "race" between these two processes: If the stop process finishes before the go process, the response is correctly withheld (signal-inhibit); but if the go processes finishes first, the response is incorrectly executed (signal-respond). Go and stop are assumed to be independent. Verbruggen and Logan (2009) reviewed the literature and concluded that the independence assumptions are met in most Stop-signal and Change-signal studies, although neural data suggest that there is a brief moment of interaction towards the end (to suppress the actual motor output; see e.g. Boucher, Palmeri, Logan, & Schall, 2007). Interestingly, the horse-race model makes no assumptions about the involvement of executive control in response inhibition - and in fact, it has been proposed that response inhibition in Stop-Signal and Change-Signal paradigms may be mediated (at least in part) by associative processes even in humans. In several studies, Verbruggen and colleagues (Verbruggen & Logan, 2008a, 2009a; Verbruggen, Best, Bowditch, Stevens, & McLaren, 2014; Best, Lawrence, Logan, McLaren, & Verbruggen, 2015) have presented evidence that a stimulus that was consistently paired

with the command to withhold a response eventually elicited automatic response inhibition. They argued that this effect could occur through associative learning.

In human cognition, executive functions are seen as crucial for the planning of actions, which demands the ability to anticipate and select appropriate responses. However, Meier et al. (2018), using pigeons, replicated at least some aspects of human performance on Stop- and Change-signal tasks. Pigeons have repeatedly demonstrated responding purely on the basis of stimulus-response associations under conditions in which humans employ more reflective processes (e.g. Lea, Wills, Leaver et al., 2009; Meier, Lea, & McLaren, 2016a, 2016b; Wasserman, Nagasaka, Castro et al., 2013; Wills, Lea, Leaver et al., 2009). Meier et al.'s results therefore support the claim that associative processes are sufficient to produce typical human behavior in such tasks, and the use of those tasks to assess executive control may therefore need re-appraisal. .

The Stop-signal and Change-signal tasks have found substantial use in experiments on rats (e.g. Feola, de Wit & Richards, 2000), and also on rhesus monkeys, using as responses both saccades (e.g. Hanes, Patterson & Schall, 1998) and arm movements (e.g. Scangos & Stuphorn 2010). In most of these studies, however, the focus has been on the neural or biochemical basis of the inhibitory responses. In addition to the experiment of Meier et al. (2018) cited above, three further experiments have taken a more purely behavioral approach to investigating the fit of the horse-race model to Stop- or Change-signal tasks in a variety of non-human species, all producing results that tend to support the model. Beuk, Beninger and Paré (2014) used rats, Knolle, McBride, Stewart et al. (2017) used sheep, and Meier, Pant, Van Horik et al. (2017) used pheasants. The tasks used in these experiments varied. Humans participating in Stop- or Change-Signal tasks are typically required to click a mouse or key in response to signals appearing on a small screen, and hence Beck et al. (2014) required their rats to press levers near signal lights, while Meier et al. (2018) required their pigeons to peck at similar stimuli on a touchscreen. However, Meier et al. (2017) and Knolle et al. (2017) deployed tasks that were quite different from those used in the human cognition

literature: they required animals to alter the course of locomotion, thus making a change response with their entire bodies. Unlike the experiments of Beck et al. (2014) and Meier et al. (2018), these experiments did not require extensive training to establish the prepotent “Go” response, but took advantage of the animals’ unconditional response of approaching food.

The present experiment was designed to bridge between these two types of procedure. It involved pigeons pecking at a touchscreen, but the stimuli were arranged so that the pigeon was in the course of an extended movement, tracking a target that moved in a regular fashion across a screen. Like the experiments of Knolle et al. (2017) and Meier et al. (2017), therefore, it created a situation in which there might be an expectation of continuity of action, which would be violated by a Change signal. By testing the predictions of the independent horse-race model in this novel situation, which links the two rather different experimental approaches that have been used so far, the experiment aimed to strengthen (or, of course, challenge) the claim that Stop- and Change-Signal tasks in general can be understood in terms of associative processes.

The simplest prediction of the independent horse-race model is that error responses made on Stop or Change trials will have a shorter mean latency than correct responses on Go trials (Logan & Cowan, 1984). This is because errors represent the case where the inhibitory process loses the horse race: the distribution of their latencies should therefore be the same as the distribution of responses on Go trials (when only the excitatory process is operative), curtailed by the distribution of latencies of the Stop or Change response. This prediction does not require any assumptions about the form or parameters of the latency distributions. Put simply, only a faster subset of the Go responses will ‘win’ the horse race on signal trials. For a simple test of the independent horse-race model, responses on a probe trial that were so fast that they prevent the occurrence of the stimulus change (referred to below as Prevention responses) should be included in the Error distribution, and this is what is commonly done in the human experimental literature (see Verbruggen et al., 2019). In the present experiment, however, they were excluded, in order to allow a comparison of

predictions with alternative models; since these are the fastest responses, excluding them conservative when testing the hypothesis that Error latencies will be relatively short.

An additional prediction of the horse-race model concerns the effect of postponing the presentation of the Stop or Change signal for a brief period (termed the Stop Signal Delay, SSD). It is obvious that, the longer the SSD, the more errors the subject will make on Stop or Change trials, because the inhibitory process is being handicapped in the horse race: it cannot start until the Stop or Change signal appears. Although this prediction is routinely confirmed in the human literature, as we noted above, standard practice in such experiments is to include trials where the subject responds to the "Go" stimulus before the Stop or Change stimulus appears as errors. For reasons explained above, we did not do that, and we therefore cannot be sure that this prediction applies in our analysis. A further quantitative prediction about Stop-Signal Delays can be made, however. Logan and Burkell (1986) found that, in a Stop-change task with humans, the latency from the moment of presentation of the change signal to the moment of successful execution of a change response was affected very little by the SSD. That is to be expected since the correct, change response is triggered by the stimulus change, so its latency timed from that moment should be the same regardless of SSD.

It might seem that it is a foregone conclusion that Error responses on Stop or Change trials should be faster, on average, than correct Go responses. It is not. It is possible that the presentation of a Stop or Change stimulus would cause either animal subjects, or human participants, to cancel any ongoing response emission process, and re-evaluate the situation as a decision between two options. Thus we can contrast the predictions of the independent horse-race model with those of models of choice reaction time, in which choice is considered as a distinct cognitive process. Such a process must involve some kind of weighing up and comparing two or more options. As Smith et al. (2018) point out, it is ordinarily assumed that choice in this sense incurs some kind of cognitive cost, which will be reflected in slower reaction times when a choice has to be made. This is seen in the

Hick-Hyman law for choice reaction time (Hick, 1952; Hyman, 1953), which though developed to describe human behavior also fits data from pigeons (e.g. Vickrey & Neuringer, 2000). So if a comparative decision process underlies responding in Stop or Change stimulus procedures, we should expect slower responding on probe trials, since on those trials a choice has to be made. The independent race model, in contrast, involves no such comparative process, and as a result it predicts faster responding to the Stop or Change stimulus on probe trials than on non-probe trials (when no choice is required; i.e. the pigeon simply has to peck the original location). However, a model involving comparison between stimuli cannot make any prediction for responses that are complete before the stimulus change occurs (i.e. before the expiry of the SSD). In order to allow a comparison between a choice-reaction model and the independent horse-race model, we excluded such "Prevention" responses from the distribution of Error response latencies.

In the present experiment, pigeons were presented with a colored target (the Go stimulus) that moved in a regular arc around a touchscreen when the pigeon pecked at it. This kind of tracking task is known to be within the pigeon's capacity, and the birds show by anticipatory responses that they have learned that movement of the target is regular (Wilkinson & Kirkpatrick, 2009). In our version of the task, on probe trials, the target changed color after a brief delay (the SSD); simultaneously, an additional target, in the normal Go color, appeared slightly towards the center of the screen from the normal trajectory. The pigeon's task was to peck the Go-colored target, wherever it appeared, causing it to move onwards. Pecking the Changed-color target counted as an error, and did not advance the target. If, on a probe trial, the pigeon pecked the Go-colored target before the SSD had been completed and the target had changed color, the scheduled change of position was cancelled. In our situation, the predictions of the horse race model for latencies in probe trials are therefore:

(1) The probability of incorrectly pecking the Changed-color stimulus will increase as the SSD increases. This prediction arises because, the longer the SSD, the more advanced the process of

responding to the stimulus in its unchanged position will be, and therefore the lower the probability that the response to the stimulus in its Changed position can beat it in the horse race.

(2) Timed from the start of the probe trial, the mean latency to peck the Changed-color stimulus (referred to below as the Error latency) will be shorter than the mean latency to peck the Go stimulus on its normal trajectory (referred to below as the Go latency). This is the key prediction, and as explained above, it arises because, under the independent horse-race model, slow Error responses are less likely to occur because they are pre-empted by successful Change responses.

(3) Timed from the moment when the target stimulus changed color (and the additional target appeared in its new position), the latency to correctly peck the target in its new position will be independent of SSD. We refer to this latency as the Change-Correct latency. It follows that timed from the start of the probe trial, the latency of a correct peck should increase with SSD. We refer to this latency below as the Start-Correct latency.

Note that it is not possible to make a prediction based on Logan and Burkell's (1986) observation that pecks to the target in its changed position will have shorter latencies for longer SSDs on probes where the subject responds erroneously. This is because error and correct responses on probe trials were not physically independent (both required the pigeon to peck), whereas in Logan and Burkell's experiment they were made with different hands.

Method

Subjects

Eight pigeons (domestic *Columba livia*) started the experiment. They were obtained as discards from local fanciers. They were maintained in an indoor aviary and transferred to cages shortly before testing. Their weights were controlled at or above 80% of their aviary free feeding weights by restricted post-test feeding, which was also done in individual cages. The pigeons are referred to here by their names, Almond, Apple, Boo, Congo, Egypt, Kar, Luna, and Pidge. Because of logistical

constraints (the retirement of the first author and consequent repurposing of the laboratory), three of the pigeons (Egypt, Kar and Luna) did not complete all the blocks of sessions described below. Four of the pigeons (Almond, Apple, Boo and Luna) had previously served as subjects in the Stop-and-Change-signal experiments of Meier et al. (2018), or in pilot work for those studies using similar procedures; the other four had served in unrelated experiments (Lea, Poser-Richet & Meier, 2015, and experiments using the same general design; Meier et al., 2016a, 2016b).

Apparatus

The experiment was conducted in four 71 x 50.5 x 43.5 cm operant chambers; each pigeon was always tested in the same chamber. One long wall of each chamber was fitted with a 31 x 23.5 cm (15-in.) touch monitor (Model 1547L 1024 x 768 pixel TFT monitor with CarrollTouch infra-red detector; ELO Touchsystems Inc.), mounted 12 cm above the grid floor of the chamber. One centimeter on the touchscreen corresponded to approximately 30 pixels. Effective pecks to target areas were followed by an immediate bleep from a 50-ohm loudspeaker, which also played white noise into the box. Two 2.8-W white houselights were mounted above and to either side of the screen. Two 6 x 5-cm apertures gave access to grain hoppers when solenoids were activated; they were located directly below the houselights and 4 cm above the floor of the chamber. The hoppers were illuminated by a 2.8-W white light when activated, and contained a 2:1 mixture of hemp seed and conditioner. The interior of some of the boxes was monitored by a video camera. The experiment was controlled by a computer (Quadvision Ltd) located in an adjacent laboratory area, using the Whisker control server system (Cardinal and Aitken, 2010) with client programs written in Microsoft® Visual Basic 6.0.

Procedure

All pigeons had previous experience in the chambers used. They were re-acclimatized to pecking the touch screen by a pretraining procedure in which a white filled circle of 80 pixels diameter (the

“observing key”) appeared centrally on the touch screen; two pecks at it led to the presentation of a random array of white filled hexagons of the same diameter. Two successive pecks at any hexagon caused it to disappear, and a white filled circle to appear close to the nearer feeder. A single peck at this circle led to the feeder being operated for 2.5 seconds. When all hexagons had been removed, there was a variable inter-trial interval of 1 to 5 seconds, followed by the reappearance of the observing key. Pretraining sessions involved eight such arrays. All pigeons but one required only a single pretraining session to ensure that they were removing all the hexagons reliably; the remaining pigeon required three sessions.

Training sessions on the tracking procedure then followed. They were divided into 36 trials, separated by a variable inter-trial interval of 2 to 4 seconds. Each trial began with the presentation of an observing key, a white filled circle of diameter 80 pixels, centered 540 pixels from the top of the touchscreen. A single peck at the observing key led to the presentation of the Go stimulus. This was a filled circle, colored green for four of the pigeons and red for the other four. In the course of a trial, it occupied six successive positions, centered on a semi-circle of radius 200 pixels, itself centered at the position of the observing key. On all odd-numbered trials, the Go stimulus was initially centered at the 4 o’clock position of this circle. Two pecks at the stimulus caused it to move to the 5 o’clock position, and so on until it reached the 9 o’clock position; two pecks at it then led to the immediate operation of the left-hand feeder for 2.5 seconds. On even-numbered trials, the Go stimulus started at the 10 o’clock position and moved clockwise in the same way until it reached the 3 o’clock position, when two pecks at it led to the immediate presentation of the right-hand feeder for 2.5 seconds. We required two pecks at each location to avoid any risk of a bird “smearing” its beak around the semicircle rather than making distinct pecks.

The median latency of the first peck made to the Go stimulus when it had moved to a new position (other than its first position, which required a different kind of movement from the others) was defined as the Go latency, and this was used in determining Stop-signal delay (SSD) values for

individual pigeons. Training on the tracking procedure was continued for each pigeon until it was reliably completing at least 30 trials per session, and median latency to do so was showing no downward trend. Across pigeons, this required from 23 to 64 sessions; details for each pigeon are given in Table 1.

Figure 1 about here

Table 1 about here

Once tracking performance was stable, probe sessions were introduced. They consisted of 41 trials. The first 5 were always training trials, using the same procedure as above. Included among the remaining 36 trials were 12 trials, randomly chosen, within which a probe occurred. Consistent with Stop-signal and Change-signal literature, probe trials occurred infrequently, so that responding to the Go stimuli remained prepotent; note that even on probe trials, no change occurred at five out of the six target positions, so the overall frequency of probes was one in 20.5 target movements. Figure 1 shows the sequence of events on probe trials. When the Go stimulus had moved to one of its positions on the trajectory (excluding the first and the last position), it subsequently moved in towards the center of the trajectory by 90 pixels, and was replaced in its normal position by a stimulus of the same size but a different color (red for pigeons whose tracking stimulus was green, and vice versa). Pecks at this stimulus (the Changed-color stimulus) counted as errors; they were recorded but had no scheduled consequences. In order to continue tracking and gain reward, the pigeon had to peck the Go stimulus in its changed position. The eight possible positions at which a probe could occur were used at random. The movement occurred after one of three SSDs, which (as in the experiment of Meier et al., 2018) were in the ratio 1:2:3 and are referred to below as Short, Medium or Long SSD. On a probe trial, if the pigeon pecked the Go stimulus before the end of the SSD, the scheduled change in position was cancelled. The actual SSD values varied between pigeons and sessions, as follows. There were four blocks of twenty sessions, within which odd sessions used training conditions in order to maintain stable performance on the basic task, and even sessions

included probe trials; thus there were ten probe sessions per block. In Blocks 1 and 2, the SSD values were 33%, 67% and 100% of the median Go latency from the first training session of Block 1. In Block 3, they were 25%, 50% and 75% of this value. In Block 4, they were 25%, 50% and 75% of the median Go latency from the training session immediately preceding the probe session concerned. All pigeons completed Blocks 1 and 2; six pigeons completed Block 3, and five pigeons completed Block 4. Table 1 includes more detail of the SSD values used.

The experimental sessions took place between May and September 2016.

Statistical analysis

Three proportions were calculated across each probe session, and their values were used for analysis. They were:

Total Error Rate: The proportion of probe trials on which the pigeons pecked the target stimulus after it had changed color, relative to the total number of probe trials

Actual Error Rate: The proportions of probe trials on which the pigeons pecked the target stimulus after it had changed color, relative to the number of probe trials in which the color change was not prevented by a peck at the target during the SSD

Prevention Rate: The proportion of probe trials in which the color change was prevented by a peck at the target during the SSD, relative to the total number of probe trials.

Latencies of all pecking responses were recorded, and their mean values within each probe session obtained. The values used for analysis were:

Go latencies on non-probe targets in the eight positions where probes could occur;

Error latencies on probe trials (i.e. latencies from the start of a probe trial to a peck at the Changed-color stimulus);

Start-Correct latencies (i.e. latencies from the start of a probe trial to a peck at the target in its changed position, without pecking the Changed-color stimulus first);

Change-Correct latencies (i.e. latencies from the moment when a target changed position to a correct peck on it, without pecking the Changed-color stimulus first).

Detailed analysis focused on the data from Block 4 since by that stage visual inspection suggested that performance was asymptotic. Only the data from the five pigeons that completed Block 4 were included in statistical analyses. The maximum number of the critical latencies available for analysis was in principle 120 (12 probe trials from each of 10 sessions) though this number was reduced by any tendency to make prevention responses (which averaged a third of all probes at the highest SSD value), and was then divided between the Change Correct and Error categories. The number of Go latencies available was much higher, coming from all the non-probe trials as well as non-probe locations on probe trials. The statistical significance of the results from Block 4 was assessed by Generalized Estimating Equations (GEE) analyses (Harding & Hilbe, 2013). In all the analyses, dependent variable values were entered as session means. The distributions of response proportions showed pronounced density at zero with a positively skewed distribution of non-zero levels, and the analysis therefore used the Tweedie distribution with a log link function. The distributions of latencies were positively skewed and the gamma distribution with a log link function was therefore used. SSD level (Short, Medium and Long) and session number were used as within-subject variables; session number was treated as a continuous variable so as to test for trends. The Total Error Rates, Actual Error Rates, and Prevention Rates were submitted to GEE analysis using as factors the SSD level (Short, Medium or Long) and the session number. The Go and Error latencies were submitted to GEE analysis using as factors the latency type (Go or Error, a within-subjects variable), the SSD level and the session number. The Start-Correct and Change-Correct latencies were submitted to separate GEE analyses using the SSD level and session number as factors. Significance tests were performed using the confidence limits of the parameters of the fitted GEE

model, rather than the overall Model Table, for greater statistical conservatism; where more than one parameter was estimated for a factor, chi-squared values for all parameters' differences from the reference value were added to give an overall significance level for the factor. There does not seem to be a currently accepted measure of effect size deriving from the output of GEE analyses, so proportions of the variance (POV) of the untransformed data accounted for by different factors and interactions were calculated from first principles following guidance in Grissom and Kim (2012) and Nakagawa and Schielzeth (2013). Overall model fit is reported in terms of the Quasi Likelihood under Independence model Criterion (QIC), in smaller-is-better form. Statistical analyses were carried out using IBM SPSS, versions 22-24. Full raw data are available at https://osf.io/9vmq3/?view_only=bffb02340d064903a92045f2838466f5

Results

As stated above, all data reported are for the five pigeons that completed all four session blocks. Results from the remaining three pigeons were qualitatively similar for the blocks that they completed. Data for non-probe conditions are reported from the test sessions only; results from the training sessions interspersed between test sessions were similar.

Descriptive statistics for data across all four session blocks

Across all four session blocks, Actual Error Rate (the mean probability of error on probe trials, given that the Change signal had appeared), was 0.18, ranging from 0.13 to 0.27 across pigeons. Actual Error Rate decreased across session blocks, from a mean of 0.21 (range 0.11 to 0.32) in Block 1 to a mean of 0.16 (range 0.06 to 0.24) in Block 4. Contrary to prediction, the pigeons showed no monotonic trend in error rates across SSD values (means of 0.16, 0.22 and 0.17 for low, medium and high SSDs). The Prevention Rate (the mean probability of responding before the Change signal could appear, when a Change trial was scheduled), was 0.23, ranging from 0.18 to 0.30 across pigeons. As

would be expected, this probability increased with SSD (means of 0.04, 0.23 and 0.44 for low, medium and high SSDs). It also varied across blocks, but not monotonically (highest value 0.30, in Block 2, lowest 0.17, in Block 4).

The mean Go latency, the time taken to move on when the tracking spot had not changed its trajectory (and was not scheduled to change it), was 2.20s (range across the five pigeons 1.21s to 3.13s). It reduced slightly across the four session blocks, from a mean of 2.43s in Block 1 (range across the five pigeons: 1.36s to 3.46s) to a mean of 2.12s in Block 4 (range 1.03s to 3.44s). It showed little variation with the SSD in use in Probe trials of the same session (means of 2.17s, 2.24s and 2.21s for low medium and high SSDs); this is to be expected since the Go latencies are drawn only from non-Probe trials.

The overall mean Error latency, the time taken to move to the tracking spot when it had changed color, was 1.96s (range across pigeons 0.80s to 3.64s). This latency reduced more substantially across the session blocks, from a mean of 2.97s in Block 1 (range 0.85s to 9.04s) to 1.37s in Block 4 (range 0.70s to 2.73s). Averaged across blocks, it was substantially shorter at low SSDs (mean 1.33s) than at medium or high SSDs (means of 2.37s and 2.33s respectively); however, this average effect was largely due to Blocks 1 and 2, and by Block 4 there was little variation in Error latency SSD (see further analysis below).

The overall mean Start-Correct latency was 3.39s (range across pigeons 1.99s to 7.18s). It reduced considerably across blocks, from a mean of 4.45 (range 2.43s to 12.40s) in Block 1 to a mean of 2.88s (range 1.52s to 5.77s in Block 4). On average, It was somewhat shorter at the low SSD level (mean 3.11s) than at medium or high SSD (means 3.66s and 3.56s respectively), and this trend was consistent across blocks.

The overall mean Change-Correct latency was 2.66s (range across pigeons 1.38s to 6.32s). It reduced across session blocks, from a mean of 3.58s in Block 1 (range 1.60s to 11.53s) to a mean of

2.28s in Block 4 (range 1.16s to 5.03s). Latencies were somewhat longer for the low and medium SSD levels (means of 2.71s and 2.85s respectively) than for the high SSD level (mean of 2.34s). This effect was most marked in Block 1; in Block 4 the trend was non-monotonic (see further analysis below).

There was no noticeable trend in any of the above values across Block 4, and little difference in them between Blocks 3 and 4. Block 4 performance was therefore judged to be asymptotic, so that the data from this block were suitable for hypothesis testing.

Hypothesis testing on data from Session Block 4

Within Block 4, the Total Error Rate varied unsystematically between SSDs, mean values being 0.12, 0.16 and 0.10 for the Short, Medium and Long SSD levels respectively. Neither the differences between these, nor the effect of session number, was significant in the GEE (QIC = 174.94). There was a significant interaction between SSD level and session number ($\chi^2_2 = 5.99, p < 0.05$), with the rate increasing across sessions by an estimated 15.6% per session at the Medium SSD level and decreasing at the Short and Long levels by an estimated 10.7% and 1.9% per session, though all the variation was erratic in all cases.

The Actual Error Rate varied somewhat more between SSDs, mean values being 0.12, 0.21 and 0.15 for the Short, Medium and Long SSD levels respectively. The differences between these were significant ($\chi^2_2 = 17.87, p < 0.001$) in the GEE (QIC = 162.65). The effect of session number, and the interaction between SSD level and session number were also significant (session number effect, $\chi^2_1 = 4.04, p < 0.05$; interaction $\chi^2_2 = 19.49, p < 0.001$). The variation of Actual Error Rate with session number was erratic at all SSD levels, but there was a slight decreasing trend overall, which was greatest (estimated value 10.7% per session) at the Low SSD level and reversed at the Medium level (estimated increase 9.8% per session).

The Prevention Rate varied sharply between SSD levels, with means of 0.01, 0.19 and 0.33 at the Short, Medium and Long SSD levels. The differences between these were significant ($\chi^2_2 = 999.01$, $p < 0.001$) in the GEE (QIC = 103.19). There was a significant, albeit again erratic, tendency for the rate to increase across sessions ($\chi^2_1 = 31.63$, $p < 0.001$), but this was entirely due to a trend at the Long SSD level (estimated increase of 12.6% per session), and the interaction between SSD and session number was correspondingly significant ($\chi^2_2 = 2550.45$, $p < 0.001$)

Figures 2 and 3 show the latency data from Block 4 relevant to the hypotheses derived from the independent horse-race model. Figure 2 shows, for each pigeon, the mean latencies of Go and Error responses (recall that only Actual Error responses contribute to the mean Error latency). The figure shows that for four out of the five pigeons, the mean Error latency was substantially shorter than the mean Go latency, as predicted by the independent horse-race model; for the remaining pigeon, there was a smaller difference in the opposite direction. In Figure 3, panel (a) shows, for each pigeon, the mean Start-Correct latency at the three SSD levels, and panel (b) shows the corresponding Change-Correct latencies. Panel (a) shows that there was some trend, albeit weak and inconsistent, for the Start-Correct latency to increase as a function of the SSD. Panel (b) shows that this trend was reduced in all five pigeons, and largely eliminated in the mean, when Change-Correct latencies were considered, consistent with the findings of Logan and Burkell (1986).

Figures 2 & 3 about here

GEE analysis confirmed that the trend shown in Figure 2 was statistically significant. In the analysis of the Go and Error latencies (QIC = 101.97), latency type had a significant main effect ($\chi^2_1 = 6.18$, $p = .013$, POV=2.84%). Session number also had a significant effect ($\chi^2_1 = 4.02$, $p = .045$, POV=0.23%), and so did its interaction with latency type ($\chi^2_1 = 4.36$, $p = .037$, POV=2.15%). The variation in latencies across session numbers was erratic, but overall there was a slight tendency towards increasing latencies (by 0.2% per session for the Go latencies, and 10.1% per session for the Error latencies). The interaction was due to a small decrease in the difference between the latency

type across sessions. In the analyses of Change latencies, no significant effects were found, either in the Start-Correct (QIC = 102.99; POVs for SSD, Session number and their interaction 2.36%, 0.10% and 0.90% respectively) or Change-Correct latencies (QIC = 127.02, POVs for SSD, Session number and their interaction 1.52%, 0.11% and 0.92% respectively).

Discussion

The tracking procedure proved to be an effective context for testing predictions of the independent horse-race model. The pigeons tracked the moving spot efficiently, and made sufficient errors when the spot changed its trajectory for stable estimates of error latencies to be obtained. However, like the operant procedure used by Meier et al. (2018), it required substantial training of the pigeons before the change signal procedure could be introduced, and in this respect it compares unfavorably with the arena methods used by Knolle et al. (2017) and Meier et al. (2017). The source of this difference is obvious. All stop or change signal procedures require a response that is prepotent; arena methods use the unconditional approach to food, whereas in the present experiment, as in that of Meier et al. (2018), we had to train a novel response up to a high level in order to make it prepotent. The corresponding advantage of our procedure is that the prepotent response is well-defined and uniform, in contrast with the relatively uncontrolled response of approaching food; and the tracking task has an intuitive appeal as ecologically relevant to many species while being easy to observe and record.

The results of our tests of the independent horse-race model are unambiguous, and consistent with previous literature. The first prediction, that the probability of incorrectly pecking the Changed-color stimulus would increase as SSD increased, appears not to have been confirmed: differences in Total Error Rate with SSD were non-significant, and the trend was non-systematic. However, in experiments with human participants, fast responses when a Change or Stop is scheduled are generally included in the error response distribution (e.g. Logan & Cowan, 1984; Verbruggen et al. 2019). We treated them as correct since the pigeons could not know that a

change was due to occur, and we feared disruption of performance if we interfered with the contingencies of reinforcement by counting such responses as errors. From the data reported, we can see what the pattern would be if Preventions were included in the Error distribution: under that definition, the mean error rate would be the sum of the Total Error Rate and Prevention Rate reported above. The mean values of this sum in Block 4 were 0.13, 0.35 and 0.48 for Short, Medium and Long SSD values, as predicted by the model and as typically found in human participants.

The second prediction, that Error responses would be made significantly faster, on average, than Go responses, was also confirmed. Our test of this hypothesis was conservative, because of our procedure of cancelling the scheduled stimulus change and treating the response as correct, if a Prevention response was made. If we had treated these responses as errors, the mean Error latencies would necessarily have been reduced (since Prevention responses necessarily had short latencies), and their difference from the Go response latencies would have been enhanced.

The third prediction, that the Change-Correct latency should be independent of SSD, was also supported. This is consistent with the findings of Logan and Burkell (1986). However, our test of this hypothesis was weak, since the tendency for Start-Correct latencies to increase with SSD was weak, and inconsistent across pigeons. In any case, we cannot lay too much emphasis on this result, since it depends on a null result, and furthermore one that comes from a small sample showing quite variable behavior.

The results support the conclusions of Meier et al. (2018), who showed that the horse-race model could predict the performance of pigeons in an operant-conditioning simulation of a stop or change signal task. Our present task was markedly different from that used by Meier et al., involving more variation in the location of the changing stimulus, and requiring the pigeon to move its head, and indeed its body, as it responded to successive Go stimuli, rather than pecking in a constant position. These differences made our task a little more like the arena tasks used by Knolle et al. (2017) and Meier et al. (2017). We conclude that, on a continuum of stop or change signal tasks, animals of a

variety of species show behavior that is consistent with the independent horse-race model. Such behavior is correspondingly inconsistent with a cognitive decision model of response determination, which as noted in our Introduction and pointed out by Smith et al. (2018), would predict longer response times on probe than non-probe trials, because of the processing cost of making a choice between two options.

Two wider implications of these results follow from this conclusion. First, our results add to a growing body of evidence, from Stop- and Change-Signal paradigms but also other types of experiment, and across a range of species, that many tasks in which animals have two response options may not be best described in the language of “choice”. Whether we are considering simulated foraging behavior (e.g. Kacelnik et al., 2011; Shapiro et al., 2008), mid-session reversal (e.g. Smith et al., 2018), or Stop- and Change-Signal paradigms (e.g. Feola et al., 2000; Knolle et al., 2017; Meier et al., 2018, and the present experiment), performance can be accurately described by assuming that each stimulus starts an independent process, and the eventual response depends only on which of those processes completes first. There is no need to invoke any process of cognitive comparison or weighing up of alternatives, of the kind that gives rise to the Hick-Hyman law in choice-reaction time experiments. In our situation, and many others, it seems that, to quote Herrnstein (1970, p. 255), “choice is nothing but behavior set into the context of other behavior”. Many of the previous papers cited here support this conclusion for essentially the same reasons as our present results, but the implication has not always been fully brought out. Furthermore, in some of the previous experiments (e.g. Meier et al., 2018), Prevention responses were not excluded from the Error distribution, so although we know that the data are consistent with the independent horsehorse-race model, we cannot be sure that they were inconsistent with a choice-reaction model.

Second, neither successful performance in a Stop- or Change-signal task, nor performance that is quantitatively consistent with the independent horse-race model, should be taken as evidence of

complex cognitive function. Although the Stop- and Change-signal tasks have been seen as tests of executive function, that overarching construct has proved hard to describe and harder still to tie to any particular behavior (Rabbitt, 1997, chap. 1); its explanatory power is therefore open to question. Our results add to the evidence that both successful Change-signal performance, and performance consistent with the independent horse-race model, can be entirely consistent with behavior being governed by associative processes, without reference to executive processes. Indeed, Verbruggen and Logan (2008) and Bowditch, Verbruggen and McLaren (2016), among others, have argued this may be the case for humans in the highly trained tasks commonly used in experiments on executive control. This should not be surprising. Acquired inhibition has been studied within the context of associative learning at least since the time of Pavlov (1927); for a recent review, see Sosa and Ramirez (2019). Only with great caution should the capacity for inhibition or stopping, as such, be taken as evidence for complex cognitive processes.

References

- Audley, R. J. (1960). A stochastic model for individual choice behavior. *Psychological Review*, *67*, 1-15. <https://doi.org/10.1037/h0046438>
- Best, M., Lawrence, N. S., Logan, G. D., McLaren, I. P. L., & Verbruggen, F. (2016). Should I stop or should I go? The role of associations and expectancies. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 115-137.
- Beuk, J., Beninger, R. J., & Paré, M. (2014). Investigating a race model account of executive control in rats with the countermanding paradigm. *Neuroscience*, *263*, 96-110. <https://doi.org/10.1016/j.neuroscience.2014.01.014>
- Bowditch, W. A., Verbruggen, F., & McLaren, I. P. L. (2016). Associatively mediated stopping: Training stimulus-specific inhibitory control. *Learning & Behavior*, *44*, 162-174. <https://doi.org/10.3758/s13420-015-0196-8>

- Feola, T. W., de Wit, H., & Richards, J. B. (2000). Effects of d-amphetamine and alcohol on a measure of behavioral inhibition in rats. *Behavioral Neuroscience*, *114*, 838-848.
<https://doi.org/10.1037/0735-7044.114.4.838>
- Grissom, R. J., & Kim, J. J. (2012). *Effect sizes for research*, 2nd edition. New York: Routledge.
- Hanes, D. P., Patterson, W. F., & Schall, J. D. (1998). Role of frontal eye fields in countermanding saccades: Visual, movement, and fixation activity. *Journal of Neurophysiology*, *79*, 817-834.
- Harding, J. W., & Hilbe, J. M. (2013). *Generalized estimating equations*, 2nd edition. Boca Raton FL: CRC Press.
- Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, *13*, 243-266. doi:10.1901/jeab.1970.13-243
- Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, *4*, 11-26. doi:10.1080/17470215208416600
- Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology*, *45*, 188-196. doi:10.1037/h0056940
- Kacelnik, A., Vasconcelos, M., Monteiro, T., & Aw, J. (2011). Darwin's "tug-of-war" vs. starlings' "horse-racing": how adaptations for sequential encounters drive simultaneous choice. *Behavioral Ecology and Sociobiology*, *65*, 547-558. doi:10.1007/s00265-010-1101-2
- Knolle, F., McBride, S. D., Stewart, J. E., Goncalves, R. P., & Morton, A. J. (2017). A stop-signal task for sheep: introduction and validation of a direct measure for the stop-signal reaction time. *Animal Cognition*, *20*, 615-626. <https://doi.org/10.1007/s10071-017-1085-7>
- Lea, S. E. G., Poser-Richet, V., & Meier, C. (2015). Pigeons can learn to make visual category discriminations using either low or high spatial frequency information. *Behavioural Processes*, *112*, 81-87. <https://doi.org/10.1016/j.beproc.2014.11.012>
- Lea, S. E. G., Wills, A. J., Leaver, L. A., Ryan, C. M. E., Bryant, C. M. L., & Millar, L. (2009). A comparative analysis of the categorization of multidimensional stimuli: II. Strategic information search in humans (*Homo sapiens*) but not in pigeons (*Columba livia*). *Journal of Comparative*

- Psychology*, 123, 406-420. <https://doi.org/10.1037/a0016851>
- Logan, G. D., & Burkell, J. (1986). dependence and independence in responding to double stimulation - a comparison of stop, change, and dual-task paradigms. *Journal of Experimental Psychology: Human Perception and Performance*, 12(4), 549–563.
- Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action - a theory of an act of control. *Psychological Review*, 91, 295-327. <https://doi.org/10.1037//0033-295X.91.3.295>
- Logan, G. D., Van Zandt, T., Verbruggen, F., & Wagenmakers, E.-J. J. (2014). On the ability to inhibit thought and action: General and special theories of an act of control. *Psychological Review*, 121, 66–95. <https://doi.org/10.1037/a0035230>
- MacLean, E. L., Hare, B., Nunn, C. L., ... and Zhao, Y. (2014). The evolution of self-control. *Proceedings of the National Academy of Sciences of the United States of America*, 111, E2140-E2148. <https://doi.org/10.1073/pnas.1323533111>
- Meier, C., Lea, S. E. G., & McLaren, I. P. L. (2016a). A stimulus-location effect in contingency-governed, but not rule-based, discrimination learning. *Journal of Experimental Psychology: Animal Learning and Cognition*, 42, 177-186. <https://doi.org/10.1037/xan0000098>
- Meier, C., Lea, S. E. G., & McLaren, I. P. L. (2016b). Task-switching in pigeons: Associative learning or executive control? *Journal of Experimental Psychology: Animal Learning and Cognition*, 42, 163-176. <https://doi.org/10.1037/xan0000100>
- Meier, C., Lea, S. E. G., & McLaren, I. P. L. (2018). Pigeons in control of their actions: Learning and performance in stop-signal and change-signal tasks. *Journal of Experimental Psychology: Animal Learning and Cognition*, 44, 82-94. <https://doi.org/10.1037/xan0000155>
- Meier, C., Pant, S. R., van Horik, J. O., Laker, P. R., Langley, E. J. G., Whiteside, M. A., Verbruggen, F., & Madden, J. R. (2017). A novel continuous inhibitory-control task: variation in individual performance by young pheasants (*Phasianus colchicus*). *Animal Cognition*, 20, 1035-1047. <https://doi.org/10.1007/s10071-017-1120-8>
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from

- generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4, 133-142.
doi:10.1111/j.2041-210x.2012.00261.x
- Pavlov, I. P. (1927). *Conditioned reflexes*. (Trans. G. V. Anrep). London: Routledge and Kegan Paul.
- Rabbitt, P. M. A. (1997). *Methodology of frontal and executive function*. Hove, UK: Psychology Press.
- Scangos, K. W., & Stuphorn, V. (2010). Medial frontal cortex motivates but does not control movement initiation in the countermanding task. *Journal of Neuroscience*, 30, 1968-1982.
<https://doi.org/10.1523/JNEUROSCI.4509-09.2010>
- Shallice T., & Burgess, P. (1993). Supervisory control of action and thought selection. In A. Baddeley & L. Weiskrantz (Eds.), *Attention: selection, awareness, and control*, pp. 171-187. Oxford: Oxford University Press.
- Shapiro, M. S., Siller, S., & Kacelnik, A. (2008). Simultaneous and sequential choice as a function of reward delay and magnitude: Normative, descriptive and process-based models tested in the European Starling (*Sturnus vulgaris*). *Journal of Experimental Psychology: Animal Behavior Processes*, 34, 75-93. <https://doi.org/10.1037/0097-7403.34.1.75>
- Smith, A. P., Zentall, T. R., & Kacelnik, A. (2018). Midsession reversal task with pigeons: Parallel processing of alternatives explains choices. *Journal of Experimental Psychology: Animal Learning and Cognition*, 44, 272-279. doi:10.1037/xan0000180
- Sosa, R., & Ramírez, M. N. (2019). Conditioned inhibition: Historical critiques and controversies in the light of recent advances. *Journal of Experimental Psychology: Animal Learning and Cognition*, 45, 17-42. doi:10.1037/xan0000193
- Verbruggen, F., Aron, A. R., Band, G. P., Beste, C., Bissett, P. G., Brockett, A. T., ... Boehler, C. N. (2019). A consensus guide to capturing the ability to inhibit actions and impulsive behaviors in the stop-signal task. *eLife*, 8, e46323. <https://doi.org/10.7554/eLife.46323>
- Verbruggen, F., Best, M., Bowditch, W. A., Stevens, T., & McLaren, I. P. L. (2014). The inhibitory control reflex. *Neuropsychologia*, 65, 263-278.
- Verbruggen, F., & Logan, G. D. (2008a). Automatic and controlled response inhibition: Associative

- learning in the go/no-go and stop-signal paradigms. *Journal of Experimental Psychology: General*, *137*, 649-672. <https://doi.org/10.1037/a0013170>
- Verbruggen, F., & Logan, G. D. (2008b). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences*, *12*, 418-424.
- Verbruggen, F., & Logan, G. D. (2009). Models of response inhibition in the stop-signal and stop-change paradigms. *Neuroscience and Biobehavioral Reviews*, *33*, 647-661. <https://doi.org/10.1016/j.neubiorev.2008.08.014>
- Vickrey, C., & Neuringer, A. (2000). Pigeon reaction time, Hick's law, and intelligence. *Psychonomic Bulletin & Review*, *7*, 284-291. doi:10.3758/BF03212983
- Wasserman, E. A., Nagasaka, Y., Castro, L., & Brzykcy, S. J. (2013). Pigeons learn virtual patterned-string problems in a computerized touch screen environment. *Animal Cognition*, *16*, 737-753. <https://doi.org/10.1007/s10071-013-0608-0>
- Wilkinson, A., & Kirkpatrick, K (2009). Visually guided capture of a moving stimulus by the pigeon (*Columba livia*). *Animal Cognition*, *12*, 127-144.
- Wills, A. J., Lea, S. E. G., Leaver, L. A., Osthaus, B., Ryan, C. M. E., Suret, M., Bryant, C. M. L., Chapman, S. J. A., & Millar, L (2009). A comparative analysis of the categorization of multidimensional stimuli: I. Unidimensional classification does not necessarily imply analytic processing; evidence from pigeons (*Columba livia*), squirrels (*Sciurus carolinensis*) and humans (*Homo sapiens*). *Journal of Comparative Psychology*, *123*, 391-405.

Table 1

Procedural details for each pigeon

Pigeon	Tracking stimulus colour	Training sessions before probes began	Blocks completed	Median Go latency from start of Probe Block 1, used to determine SSDs in Blocks 1-3 (ms)	Median Go latency from preceding training sessions, used to determine SSDs in Block 4 (ms)	
					Minimum	Maximum
Almond	green	36	1-4	1369	1345	2104
Apple	red	25	1-4	1535	1055	1530
Boo	red	25	1-4	1274	974	1362
Congo	green	23	1-4	1816	1312	1948
Egypt	green	64	1-2	1166		
Kar	red	25	1-2	2007		
Luna	green	54	1-3	1731		
Pidge	red	25	1-4	872	694	902

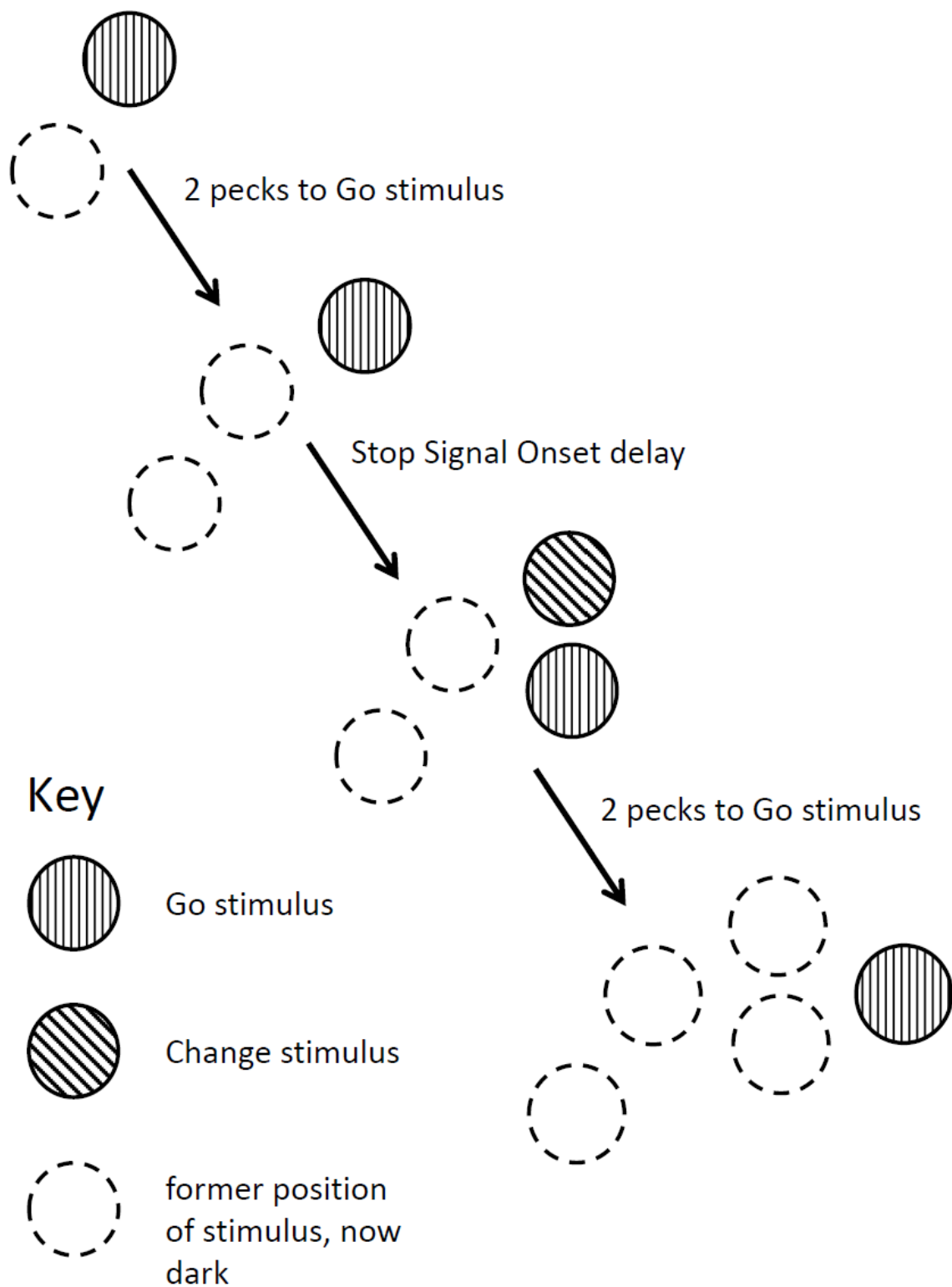


Figure 1: Partial sequence of events within a probe trial. Note that the pecks to the Changed-colour stimulus had no scheduled consequences, and that pecks to the Go stimulus before the Stop Signal Delay had elapsed cancelled the probe. Not to scale.

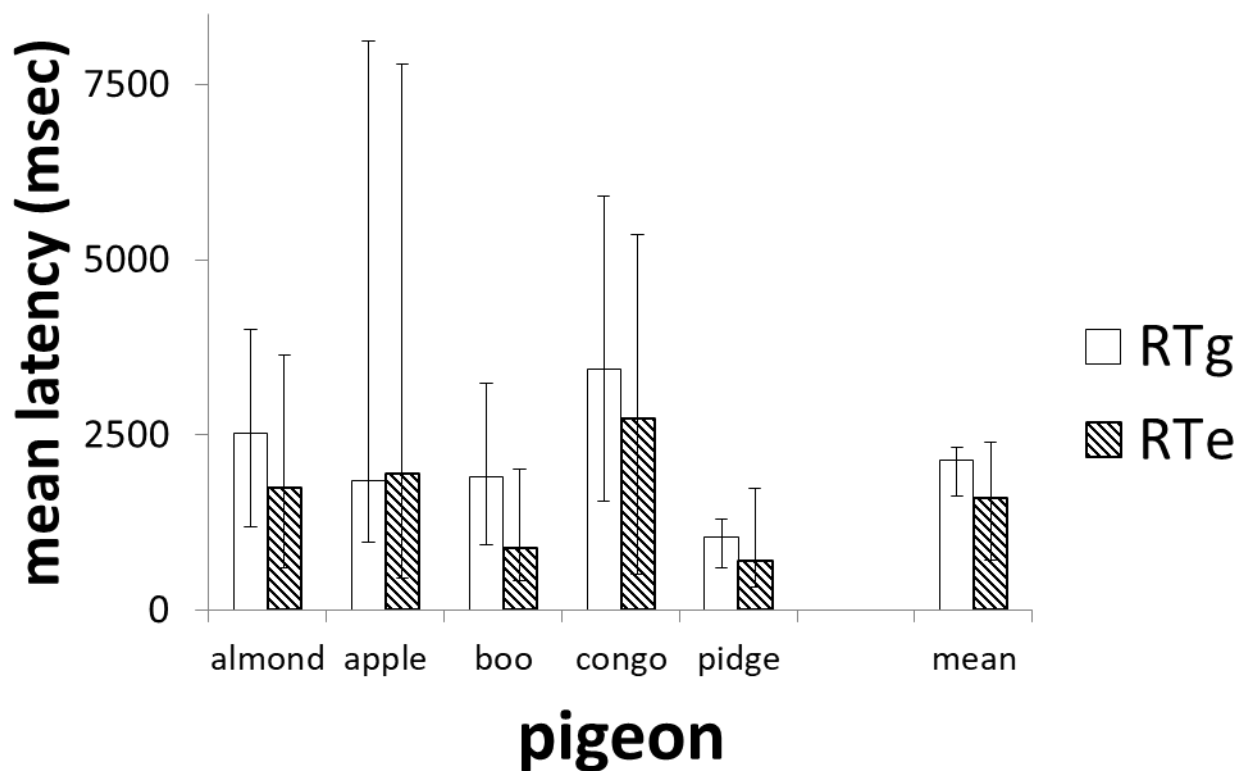
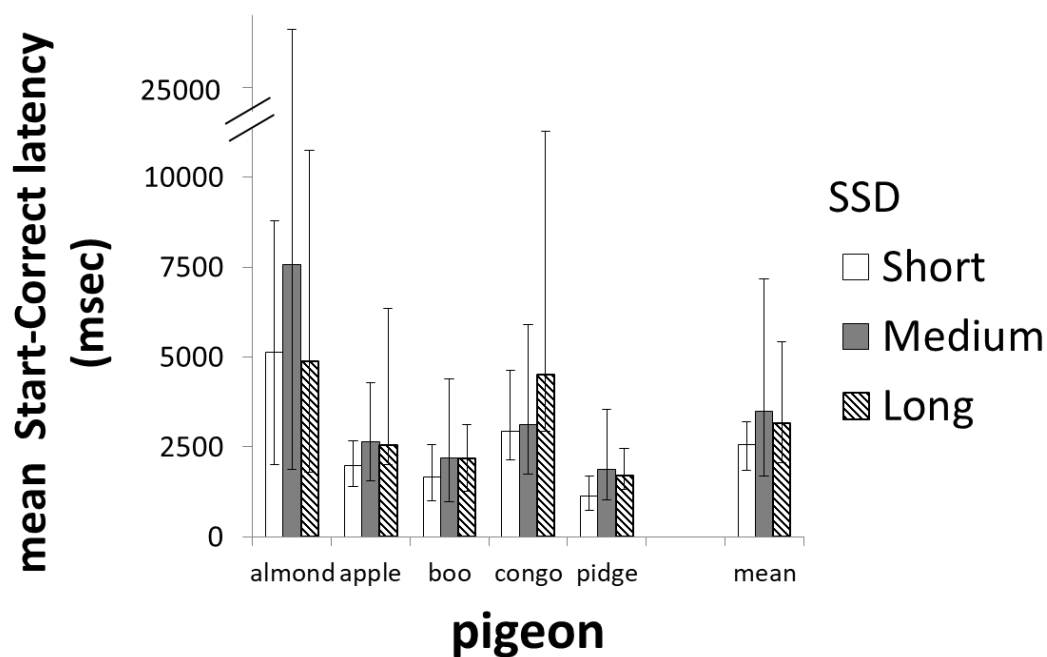


Figure 2: Mean latencies of Go responses (RTg) and Error responses (RTe) in Test Block 4.

Error bars show ranges of these means over the 10 sessions of the block.

(a)



(b)

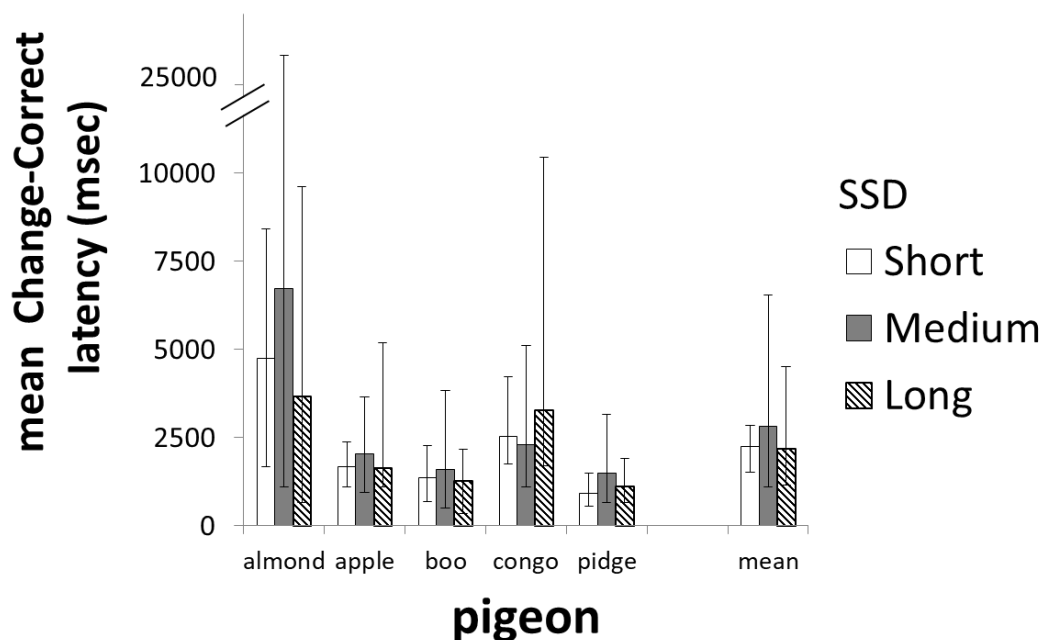


Figure 3: Mean latencies of Change responses as a function of Stop Signal Delay (SSD), in Test Block 4. (a) Start-Correct latencies (timed from the appearance of the Go stimulus in its unchanged position) and (b) Change-Correct latencies (timed from appearance of the Changed-colour stimulus). Error bars show ranges of these means over the 10 sessions of the block.