LIBRARY
Te Tumu Herenga
THE UNIVERSITY OF AUCKLAND

RESEARCHSPACE@AUCKLAND

http://researchspace.auckland.ac.nz

ResearchSpace@Auckland

Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognise the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

To request permissions please use the Feedback form on our webpage.
http://researchspace.auckland.ac.nz/feedback

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form and Deposit Licence.

Note : Masters Theses

The digital copy of a masters thesis is as submitted for examination and contains no corrections. The print copy, usually available in the University Library, may contain corrections made by hand, which have been requested by the supervisor.

The Discriminative Functions of Primary and Conditional Reinforcers:

Signalling the Local and Global Contingencies of Reinforcement

Nathalie Boutros

A thesis submitted in partial fulfilment of the requirement for the degree of Doctor of Philosophy

Department of Psychology

The University of Auckland

2010

Abstract

In 4 experiments, pigeons worked in two-key concurrent schedules for food and non-food response-contingent events. Choice after one of these events was a function of the global and local relative probability of a same-alternative food. Experiment 1 was a steady-state two-alternative concurrent-schedule procedure with added response-contingent red keylights, whose left: right ratio was positively, negatively or zero-correlated with the left: right food ratio. Local preference after a red keylight was always towards the just-productive alternative, regardless of the stimulus ratio-food ratio correlation. Pairing the stimuli with food enhanced this effect. In Experiment 2, response-contingent keylights signalled the likely location ($p = .9$) of the next food, and preference was towards the locally richer alternative, whether this alternative was the just-reinforced or the not-just-reinforced alternative. When the two alternatives were equally likely to produce the next reinforcer, preference was towards the just-reinforced alternative. This was because the post-event changeover contingencies biased the local obtained food ratio. This was confirmed in Experiment 3 in which the post-food illuminated alternative was varied and food was the only response-contingent event. Local preference was always towards the post-food illuminated alternative when the reinforcers randomly alternated. When the reinforcers strictly alternated, preference was initially towards the post-food illuminated alternative before changing to the not-just-reinforced alternative. This finding confirmed that previous difficulties with strict-alternation were likely due to the post-food changeover contingencies biasing the perceived post-food obtained local reinforcer ratio. Experiment 3 also revealed that preference was shifted by same-alternative reinforcers (continuations) regardless of the post-food changeover contingencies, suggesting a response-strengthening function of temporally distant reinforcers. Experiment 4 revealed that control by temporally distant reinforcers is apparently not discriminative: there was no control by the local probability of a same-alternative reinforcer (*p* = 0 or 1) when sequences of same-alternative reinforcers strictly alternated. Preference was instead a function of the global probability of a continuation reinforcer. Together, these experiments demonstrate that response-contingent stimuli (appetitive and non-appetitive) function as signals indicating the likely location of subsequent appetitive stimuli. They can signal the short-term, or the long-term contingencies of further appetitive stimuli, or both.

Acknowledgements

I would like to first thank my supervisors, Professor Michael Davison and Dr. Douglas Elliffe for the use of their equipment as well as for all of their expertise, advice, corrections, experience and support. Moreover, I thank them for allowing me the freedom to explore and to learn. This thesis never would have been completed without them. I also thank Mr. Mick Sibley for his excellent care of the subjects. I would also like to thank all of the students that have come through the operant lab in my time here. Thanks for helping to care for the animals, and thanks for some very interesting conversations. Lastly, I thank Neil Green for all of his support and care.

Table of Contents

List of Tables

Page

Chapter III

Chapter IV

Chapter V

Chapter VI

List of Figures

Page

Chapter III

Chapter IV

Chapter V

Chapter VI

Chapter I

A conditioned or conditional reinforcer (Dinsmoor, 2004) is a stimulus, initially hedonically neutral which, through some relationship with an inherently appetitive stimulus (a primary reinforcer, e.g., food), comes to have effects on behaviour similar to those of those of the primary reinforcer to which it is related. Traditionally, a conditional reinforcer is created by presenting the initially neutral stimulus either at the same time as, or immediately prior to, the hedonic one (Kelleher & Gollub, 1962). After many of these pairings, and few or no presentations of the appetitive and non-appetitive stimuli in isolation, the non-appetitive stimulus becomes a conditional reinforcer, that is, able to itself maintain responding. This is a general description of a procedure with many variants. For example, no response may be required to produce either the conditional reinforcer or the appetitive stimulus (e.g.,in autoshaping; Brown & Jenkins, 1968); or a response may be required only to produce the conditional reinforcer (e.g., a delay of reinforcement procedure; Schaal & Branch, 1988); or both the conditional and primary reinforcers may be response-contingent (e.g., concurrent chains procedures; Herrnstein, 1964).

How should conditional reinforcers be characterized? Williams (1991b; 1994) argued that conditional-reinforcer effects are due to the initially hedonically neutral stimulus acquiring some of the properties of the primary reinforcer with which it has been paired (including the appetitive properties). This "conditioned value" account holds that a stimulus paired with a primary reinforcer itself *becomes* a reinforcer through some process of Pavlovian conditioning: when the primary reinforcer and the initially neutral stimulus are repeatedly paired, value is transferred from the former to the latter. This is the most widely-held view of conditional reinforcement and practical applications reflect this. For example, there have been a number of attempts to increase consumption of less preferred but healthful foods by pairing these foods with highly preferred items or activities (Ahearn, 2003; Kern & Marder, 1996; Piazza et al., 2002; Riordan, Iwata, Finney, Wohl, & Stanley, 1984; Solberg, Hanley, Layer, & Ingvarsson, 2007; Tiger & Hanley, 2006). Increased consumption subsequent to pairing is said to occur because of the food's greater value, brought about by a Pavlovian process (Piazza et al., 2002).

In perhaps the least-confounded assessment of the causes of increased food acceptance after pairing with a preferred stimulus, Solberg et al. (2007) found that the increased preference for the initially less preferred food did not persist when pairings were eliminated (see also Ahearn, 2003). Solberg et al. concluded that the conditioned value acquired by the initially less preferred food was transient. This conclusion, along with the general assumption that conditional reinforcement involves a transferral of value from the hedonic to the neutral stimulus, arises from very dated Pavlovian conditioning theory and research.

1.1 Pavlovian conditioning

According to Pavlov (1927), creating a conditional stimulus (CS), whereby an initially neutral stimulus comes to elicit a reflex, requires that the stimulus be repeatedly presented in close temporal contiguity with a stimulus which does so unconditionally. When an unconditional stimulus (US) is paired with an initially neutral stimulus, a path is said to be created between the neutral stimulus and the same point in the central nervous system to which the US connects (although the connection may not be as strong). Thus, the conditional stimulus comes to elicit the same responses as the unconditional stimulus because it, in effect (although perhaps only weakly), substitutes for the unconditional stimulus. Pavlov only spoke in terms of (hypothesized) physiological processes and was in fact admittedly wary of himself and his colleagues “becoming psychologists”, thus expressly avoiding any discussion of mental states (Pavlov, 1904). Despite this, the conditioned value account in which conditional reinforcers are said to exert their effects because of their acquired hedonic properties (e.g., Williams, 2003) can be seen as having their roots in Pavlov’s stimulus-substitution characterisation.

More recent theories of Pavlovian conditioning have been within a framework emphasizing correlations between events, and the informational properties of the stimulus paired with a primary reinforcer (Balsam, Fairhurst, & Gallistel, 2006; Egger & Miller, 1962, 1963; Rescorla, 1967, 1972, 1988). According to these theories, an initially-neutral stimulus will only become a CS if that stimulus provides information about the arrival of the US. Simply presenting a neutral stimulus at the same time as, or slightly before, food on a number of trials is no guarantee that the stimulus will become a CS. The information-centric account does not require that the CS acquire any of the properties of the US, including the hedonic or appetitive ones. What is required is that the stimulus predict the arrival of the US.

In an early experiment, rats and rabbits were presented with two element stimulus compounds that were sometimes followed by food (Wagner, Logan, & Price, 1968): Group 1 was presented with stimulus compound AX+ (Stimulus A then Stimulus X then food) and on separate trials, was presented with stimulus compound BX- (Stimulus B followed by Stimulus X and the absence of food). Group 2 was presented with the same AX and BX compounds but food was equally likely after each. Although Stimulus X was, for both groups, more closely paired with food than either Stimulus A or B, Stimulus A was a better predictor of food for Group 1.

Wagner et al. (1968) reported more conditioning of Stimulus X in Group 2. If conditioning was simply a matter of the temporal contiguity of primary and conditional reinforcers, Stimulus X should have been conditioned to the same degree in both groups. Instead, the degree of conditioning of Stimulus X was greater when there was no other, more reliable source of information about the arrival of food. This *relative validity of effect* has been demonstrated in pigeons (Wasserman, 1974) and humans (Wasserman, 1990) as well as the rats and rabbits in Wagner et al.'s experiments.

Correlation can be sub-divided into necessity and sufficiency (Wasserman & Miller, 1997). Sufficiency is the probability of the US given the CS, $p(\text{US}|\text{CS})$, and necessity is the probability of the US given no CS, $p(\text{US}|\text{no CS})$. As the former probability approaches 1 and the latter approaches 0, correlation increases and the CR is predicted to increase (Rescorla, 1967). Rescorla (1968) not only found no conditional response when $p(\text{US}|\text{CS}) = p(\text{US}|\text{no CS})$, but that as $p(\text{US}|\text{CS})$ increased and $p(\text{US}|\text{no CS})$ decreased, conditioning increased, even as the absolute number of CS-US pairings remained unchanged. Only a focus on contingency, as opposed to contiguity, can account for this.

Such a focus on contingency allows for a parsimonious understanding of not only the conditioning that occurs when $p(\text{US}|\text{CS}) > p(\text{US}|\text{no CS})$, and the lack of conditioning when $p(\text{US}|\text{CS}) = p(\text{US}|\text{no CS})$, but also the inhibitory conditioning that is obtained when $p(\text{US}|\text{no CS}) > p(\text{US}|\text{CS})$. Working within a fear-conditioning paradigm, Rescorla (1966) arranged presentations of a fear-inducing US and a neutral CS in three groups. In Group 1 $p(\text{US}|\text{CS}) = p(\text{US}|\text{no CS})$, in Group 2 $p(\text{US}|\text{CS}) > 0$ and $p(\text{US}|\text{no CS}) = 0$, and in Group 3 $p(\text{US}|\text{CS}) = 0$ and $p(\text{US}|\text{no CS}) > 0$. The CS elicited a fear response when it reliably preceded shock (Group 2), had no effect when the probability of shock was the same in its presence and absence (Group 1), and inhibited the fear response when it reliably preceded an absence of shock (Group 3). Thus, the behavioural effects of the CS were a function of the correlation (positive, zero, or negative) between CS and US presentations. Moreover, prior experience with completely uncorrelated presentations of the CS and US, such as was arranged for Rescorla's Group 1, retards learning if a positive correlation is later introduced (Mackintosh, 1973). This *learned irrelevance* is not due to prior exposure to the CS, the US, or the summation of these effects. Rather, acquisition of the appropriate response is hindered because the CS had already been conditioned as irrelevant by the time the positive correlation is introduced. That there can be clear learning about the relationship between the CS and the US even when the stimuli are systematically unrelated clearly indicates that temporal contiguity is not a requirement for Pavlovian conditioning. Pavlovian conditioning instead appears to involve learning the degree to which the hedonically neutral stimulus predicts the biologically relevant one.

1.2 The information hypothesis: predictive stimuli in an operant context

Can biologically neutral stimuli which mimic the effects of primary reinforcers in operant conditioning procedures perhaps also be understood as primarily informative rather than primarily hedonic? One alternative to the pairing account is Hendry's (1969) information hypothesis, according to which a stimulus is not reinforcing by virtue of temporal contiguity with a primary reinforcer, but because it reduces uncertainty about the availability of primary reinforcement. The major evidence in favour of this characterization of conditional reinforcement is the preference for, or higher response rate in, situations where the current probability of reinforcement is signalled over situations with equivalent contingencies of reinforcement but no signals indicating the current reinforcer probability (Lewis, Lewin, Muehleisen, & Stoyak, 1974; Perone & Baron, 1980; Perone & Kaminski, 1992; Schrier, Thompson, & Spector, 1980; Wilton & Clements, 1971). According to the information hypothesis, uncertainty is aversive and any stimulus which removes the animal from this unpleasant state of uncertainty will acquire hedonic value. Thus, if periods of reinforcement randomly and unpredictably alternate with periods of extinction, not only will a stimulus which signals a period of reinforcement (S+) acquire value, but so too will a stimulus which signals a period of extinction (S-). Both stimuli equally reduce uncertainty about the arrival of primary reinforcement. Thus, both stimuli should acquire value as conditional reinforcers according to the information hypothesis. When the S+ stimuli are removed in such observing response procedures, there is generally a large reduction in responding. The reduction in response rate is generally much smaller when the S- stimuli are removed (Case, Ploog, & Fantino, 1990; Dinsmoor, Browne, & Lawrence, 1972; Gaynor & Shull, 2002; Jenkins & Boakes, 1973; Katz, 1976; Mueller & Dinsmoor, 1984; Wald & Dukich, 1978) . Thus, the S- stimuli are not as effective in maintaining responding as are the S+ stimuli. This suggests that while the S+ stimuli do have some acquired appetitive properties, the S- stimuli do not (at least to the same degree as the S+ stimuli). Some evidence even suggests that the S- has some acquired aversive properties (Blanchard, 1975; Gaynor & Shull, 2002; Purdy & Peel, 1988). Thus, although information provided about the arrival of primary reinforcement may be crucial to the behavioural effects of a conditional reinforcer (as demonstrated by the above evidence from Pavlovian conditioning procedures), this information does not appear to endow the stimulus with value as a conditional reinforcer. While the information hypothesis does depart from traditional accounts in that it does not require close temporal contiguity of the primary and conditional reinforcers, it nonetheless remains a conditioned-value account.

1.3 An alternative to conditioned-value accounts of conditional reinforcement

Does the above discussion suggest that information is irrelevant to conditional reinforcement in operant procedures? The Pavlovian literature suggested a central role for information about the arrival of primary reinforcers in generating conditional reinforcer effects. Do different processes operate in Pavlovian and operant conditioning procedures, with the discriminative properties of the stimuli primarily controlling responding in Pavlovian procedures, and the hedonic properties controlling behaviour in operant procedures? Such a distinction is plainly absurd as stimuli claimed to acquire conditioned value in operant procedures are said to do so via Pavlovian processes. In order for research within operant conditioning to remain consistent with respondent theory, the effects of conditional reinforcers must also be understood as arising from the informational properties of the stimuli. However, the above-cited research demonstrating that stimuli which signal the absence of reinforcement are not themselves reinforcers suggests that conditional reinforcer effects may not be due to any acquired *value* of the stimuli.

1.3.1 Observing responses

Evidence against a characterization of conditional reinforcer effects as arising from some acquired hedonic or appetitive properties of the stimulus has recently been obtained in operant procedures by Shahan and colleagues (Podlesnik, Jimenez-Gomez, Ward, & Shahan, 2009; Shahan, Magee, & Dobberstein, 2003; Shahan & Podlesnik, 2005, 2008a, 2008b). The general procedure of this series of experiments is as follows: A multiple schedule is arranged with components consisting of mixed schedules where a variable interval (VI)-food schedule randomly and unpredictable alternates with an extinction (EXT) schedule. No discriminative stimuli indicate whether the VI or EXT schedule is currently active. Responses to a second, observing, key do not produce food or change the schedule of food reinforcement but intermittently produce discriminative stimuli indicating whether a VI component or an EXT component is currently in effect (an observing response procedure). The (S+) discriminative stimuli produced on the observing key are often considered conditional reinforcers (Dinsmoor, 1985).

Shahan and colleagues have varied the rate of primary reinforcers (Shahan, Magee, & Dobberstein, 2003), the rate of observing response stimuli (putative conditional reinforcers; Shahan & Podlesnik, 2005), and the "value" of the conditional reinforcers (Shahan & Podlesnik, 2008a). When conditional-reinforcer rate or value is varied across components, the obtained baseline response rate is consistent with a characterization of the observing stimuli as reinforcers: more responding is obtained with higher rates or more highly valued stimuli (Herrnstein, 1970). However, tests of resistance-to-change (presentation of supplementary feed or extinction of responses previously followed by primary and conditional reinforcers) suggest that the observing stimuli do not in fact increase response strength. Typically, responding maintained with a higher rate of primary reinforcement is more resistant to change (Nevin & Grace, 2000). Responding maintained with a higher rate of observing response stimuli or by putatively more highly valued stimuli is no more resistant to change than responding maintained by lower rates or less valued stimuli. In fact, lower valued conditional reinforcers may even increase resistance-to-change. These findings have led these authors to suggest that conditional reinforcers may increase response rate through some mechanism other than increasing response strength (Shahan & Podlesnik, 2008b).

1.3.2 Second order schedules

Conditional reinforcers may have their response-increasing effects by signalling the contingencies of (primary) reinforcement. A large body of research using second-order schedules is consistent with this. In a second-order schedule, a series of simple (e.g., lever-press or key peck) responses on a first-order schedule is defined as the operant and is reinforced according to another (second-order) schedule (Marr, 1979). In a second-order schedule of the form FR 4 (FI 20-s) four FI 20-s components must be completed before primary reinforcement becomes available. A stimulus can be presented upon completion of every component, and thus is paired with food in the last component, putatively becoming a conditional reinforcer (Kelleher, 1966). Such brief stimuli produce effects similar to those produced by unconditional reinforcers: appropriate patterning of responding (e.g., a post-stimulus scallop pattern when brief stimuli are presented on a fixed-interval schedule) and changes in response rate typically produced by food (Gollub, 1977).

However, the stimuli do not have to be paired with unconditional reinforcers in order to exert such effects (Stubbs, 1971; Stubbs & Cohen, 1972). Stimuli presented upon completion of all components *except* the last one also produce effects similar to those produced by paired stimuli and by unconditional reinforcers. Such stimuli cannot be conditional reinforcers according to accounts which stress the importance of local pairing (temporal contiguity) with an unconditional reinforcer (e.g., Kelleher & Gollub, 1962).

The stimuli in these procedures may instead have their effects by signalling the behaviour which produces food (Neuringer & Chung, 1967; Reed & Hall, 1989). In a *percent reinforcement* procedure (Ferster & Skinner, 1957), non-food stimuli replace some food deliveries in a standard schedule of reinforcement and produce the same effects because, according to Neuringer and Chung, the responses that produce food also produce the stimuli. The putative conditional reinforcer is not assumed to have any acquired hedonic value; conditional reinforcer effects are simply due to the stimulus signalling the response which leads to food.

1.4 Conditional reinforcers as informative: a non-value transfer account

Although a stimulus which has been systematically related to a reinforcer may change behaviour, this behaviour change may not be due to any appetitive or hedonic value acquired by the stimulus. Conditional reinforcer effects may instead be due to the stimulus signalling something about the arrival of primary reinforcement (Davison & Baum 2006; 2010; Shahan 2010). This should not be taken to suggest that the stimulus necessarily has no acquired appetitive or hedonic properties; rather, that the motivational properties (if any) are not what is responsible for the behaviour before, during or after the stimulus. This behaviour is better considered a function of the discriminative properties of the conditional reinforcer, arising because of what the stimulus signals about the forthcoming contingencies of (primary) reinforcement.

This characterization of conditional reinforcer effects has the benefit of easily and parsimoniously accounting for conditioned inhibition (Savastano, Cole, Barnet, & Miller, 1999). If the stimulus signals that the preceding response is unlikely to be followed by a primary reinforcer, conditional reinforcer effects would not be predicted (contrary to predictions of Hendry's information hypothesis). A conditioned inhibition effect, in which there is a relative decrease in the conditioned response, is likely to be obtained. These conditioned inhibition effects are not due to any acquired aversive properties of the stimulus. Instead, the effects of the stimulus are due to what it signals about primary reinforcement, in this case that is unlikely.

All behaviour is choice behaviour (Herrnstein, 1970) and responses are emitted in direct proportion to the relative reinforcement received. A clear signal that one response is highly likely to be reinforced will lead to an increase in that response, while a clear signal that a response is highly unlikely to be reinforced will lead to an increase in all other responses (and thus a relative decrease in the target response). No concept of acquired value is required: the animal simply emits the response more likely to be followed by a reinforcer. In standard observing response procedures, response-contingent discriminative stimuli signal whether the current schedule is one of food reinforcement or extinction. As discussed above, observing responses are maintained when only the S+ stimulus is delivered (when there is no stimulus change during the extinction component). When observing responses are only ever followed by the S- stimulus, such responding is not maintained. However, there is an exception to this: when the target response is arduous, or when there are other (experimentally defined) responses available in the situation, responding for the S- alone is maintained (Case, Ploog, & Fantino, 1990; Lieberman, Cathro, Nichol, & Watson, 1997; Perone & Baron, 1980; Perone & Kaminski, 1992). In this case, the stimulus signals which response (the alternative rather than the target response) is momentarily more likely to be reinforced. Responding for the S- is also maintained when responding during the extinction component delays food in the food component (Allen & Lattal, 1989). In this case, the S- signals that there are (delayed) consequences to emitting the target response. Thus, when the S- is informative, in the sense of signalling something about the effect of current behaviour on either immediate or delayed reinforcers, responding occasionally followed by the S- will be maintained. This increase in responding should not be attributed to the motivational properties of the stimulus: even responding followed by noxious stimuli (e.g., electric shock; loud tone) can be maintained when the stimulus signals a period of reinforcement (Ayllon & Azrin, 1966; Azrin & Holz, 1966; Holz & Azrin, 1961; McMillan & Morse, 1967; Stubbs & Silverman, 1972). The crucial determinant to the behavioural effects of the stimulus is not its status as appetitive or aversive but rather its properties as a signal of further biologically relevant stimuli.

This characterization of conditional reinforcement has much in common with that proposed by Shahan (2010). Shahan first reviewed some of the literature on conditional reinforcement, with emphasis on his and his colleagues work demonstrating that conditional reinforcers, unlike primary reinforcers, do not increase resistance-to-change. Shahan then suggested that conditional reinforcers may be better understood as signposts, guiding the animal towards primary reinforcement. According to this account, stimuli identified as conditional reinforcers have their effects on behaviour not because of any acquired motivational properties, but because they signal the contingencies of reinforcement. Reinterpreting previously reported results as due to the conditional reinforcer's role as a discriminative stimulus does not require the abandonment of any of the models that have successfully described conditional reinforcer effects (e.g., Mazur, 2001; Squires & Fantino, 1971). While the interpretation of the results may differ, a characterization focusing on signalling rather than acquired value does not necessarily predict divergent findings.

1.4.1 Davison and Baum's (2006) experiment

Davison and Baum's (2006) finding is central to the present characterization of conditional reinforcer effects as primarily discriminative rather than as primarily strengthening. Davison and Baum inserted response-contingent magazine illuminations unaccompanied by food into a frequently-changing (Davison & Baum, 2000) concurrent schedule of response-contingent foods. According to a pairing hypothesis or conditioned value account, magazine-light illuminations should be conditional reinforcers for these pigeon subjects with long experimental histories of magazine light-food pairings. In their Experiment 1, Davison and Baum (2006) arranged foods (accompanied by magazine illumination) as well as magazine-lights presented without food, contingent on key-pecks. They varied the relative probability of an event being a magazine-light alone (vs. food + magazine-light). According to a value-transfer account of conditional reinforcement, as the proportion of magazine-lights presented without food increased, the value of the magazine-light as a conditional reinforcer should decrease (e.g., Dunn, Williams, & Royalty, 1987). Further, Davison and Baum (2006) reasoned that, according to a conservation of value principle, the relative value of primary reinforcers should increase as less of their (finite) appetitive value is transferred to the conditional reinforcer.

Davison and Baum (2006) judged the status of the magazine light as a conditional reinforcer by comparing the *preference pulse* (log response ratio as a function of time elapsed since the event; Davison & Baum, 2002) after the magazine light with the preference pulse after a food reinforcer. A food reinforcer typically produces a transient increase in preference to the alternative that delivered the last reinforcer. Choice then typically descends to a level representative of the overall reinforcer ratio (Davison & Baum, 2002; Landon, Davison, & Elliffe, 2003a). The initial period of relatively extreme preference to the just-productive alternative can be, and has been interpreted as a demonstration of the law of effect: an increase in the response that immediately preceded the last reinforcer (Kelleher & Gollub, 1962; Skinner, 1938). Preference pulses can thus be used to evaluate the status of a stimulus as a conditional reinforcer. Davison and Baum (2006) however found no change in the preference pulse as the proportion of magazine-lights accompanied by food decreased, suggesting that devaluing magazine-light as a conditional reinforcer had no effect on the status of this stimulus as a reinforcer.

To account for this perplexing finding, Davison and Baum (2006) noted that although increasing the proportion of magazine lights unaccompanied by food degraded the value of magazine-lights in one sense, in another sense the value of the magazine-light stimuli was unchanged. Changing the proportion of magazine-lights not followed by food had no effect on the left:right ratio of either magazine lights or foods. The correlation of the relative magazine-light rate on one alternative with the relative food rate on that alternative remained constant (at +1.0) regardless of how many magazine lights were accompanied by food.

In the frequently-changing procedure, each reinforcer signals the current component reinforcer ratio in the sense that the richer alternative in the concurrent schedule delivers more reinforcers. Thus, the just-productive alternative is more likely to be the richer alternative. In Davison and Baum's (2006) Experiment 1, not only did foods signal the food ratio, but so too did the magazine lights. The alternative that delivered the greater number of magazine lights also delivered the greater number of foods and increasing the proportion of magazine lights presented without food had no effect on this.

In their Experiment 2, Davison and Baum (2006) further investigated the role of the correlation between the relative number of non-food stimuli on an alternative with the relative number of foods on that alternative. Magazine lights could either be positively, negatively, or zero correlated with food. When the correlation was positive, the alternative that delivered the greater number of magazine lights also delivered the greater number of foods (as in their Experiment 1); when the correlation was negative, the alternative that delivered the lesser number of magazine lights delivered the greater number of foods; when the correlation was 0, both alternatives delivered equal numbers of magazine lights regardless of which alternative delivered the greater number of foods. A transient increase in preference to the just-productive alternative only followed magazine lights when they were positively correlated with food. This was also true when green keylights, which unlike magazine lights had no history of being paired with food, were used in place of magazine lights: local preference after a stimulus was a function of the food-ratio:stimulus-ratio correlation, and not of the pairing history of the stimulus with food, arguing against a conditioned-value account and in favour of a signalling account. Preference after a magazine-light was towards the alternative signalled more likely to deliver the next food, and not towards the alternative that provided that last magazine-light. In fact, Davison and Baum found a *decrease* in preference to the just-productive alternative immediately after a stimulus negatively correlated with food, even when that stimulus was a magazine light with its extensive history of pairing with food. These results were later replicated by Davison and Baum (2010).

1.5 The importance of the novelty of information

Although the CS-US correlation is important to theories of classical conditioning, a positive correlation with a US is not sufficient for the neutral stimulus to become a CS (Papini & Bitterman, 1990). For example, when two stimuli, one more salient than the other, are together paired with a US, *overshadowing* (Pavlov, 1927) is likely to occur — the stimulus with the lesser salience acquires less ability to elicit the CR compared to control procedures where it is presented in isolation. The relative probability of the US given the CS is the same in control and overshadowing preparations, yet conditioning to the less salient CS differs. The uniqueness or non-redundancy of the information provided by the CS about the arrival of a US must be considered in order to explain this and other similar findings. For example, Egger and Miller (1962) arranged that after the start of a short stimulus S1, an even shorter stimulus S2 would be presented and then followed by food. S2 acquired a greater ability to maintain responding in subjects for whom it was uniquely informative of the arrival of food (subjects who received S1 presentations not followed by S2 or food in addition to presentations of S1-S2-food). However when the information provided by S2 was redundant, more responding was maintained by contingent presentations of S1. Although S2 was more closely paired in time with primary reinforcement, the information provided about primary reinforcement was redundant with the information already present (from S1).

In the influential Rescorla-Wagner (1972) model, the change in associative strength of a CS on a trial is the difference between the maximum associative strength supportable by the US and the current total associative strength for all stimuli paired with that US (weighted by the salience of the CS and US):

$$\Delta V_x^{n+1} = \alpha_x \beta\left(\lambda - V_{tot}\right) \qquad \text{(Equation1.1a)}$$

$$V_x^{n+1} = V_x^n + \Delta V_x^{n+1}$$ (Equation 1.1b)

where ΔV_x is the change in associative strength of the Stimulus *X* (the CS), α is the salience of Stimulus *X* (between 0 and 1), β is a rate parameter which may be taken as the "associability" of the US and is closely related to its intensity (between 0 and 1), λ is the maximum associative strength supportable by the US, V_x is the current associative strength of Stimulus *X* and V_{tot} is the total associative strength of all CSs paired with the US. In Equation 1b, V_x^{n+1} is the associative strength of Stimulus *X* after trial *n*+1. The change in associative strength to Stimulus *X* on trial *n*+1 is summed with the associative strength of Stimulus *X* prior to that trial. The principle improvement of this model over its immediate predecessor (Bush & Mosteller, 1951) is that the total associative strength supportable by the US, λ, is shared amongst all CS's, rather than learning occurring to each of them independently,.

The parenthetical term in Equation 1a is said to represent the degree to which the US was unexpected on a trial and its movement towards zero as a reduction in "surprisingness" of the US (Miller, Barnet, & Grahame, 1995). Learning is said to proceed when the difference between the obtained and expected US on a trial is nonzero. After a number of trials, the US is no longer surprising after a CS ($V_{tot} \approx \lambda$) and there is little learning. This necessity of a surprising, or unexpected US accounts for various cue-competition phenomenon such as the overshadowing described above as well as *blocking* (Kamin, 1969). When a CS already adequately predicts the arrival of a US, that US is not surprising when it arrives and there is no learning of the association between the US and a second CS. If the US is made surprising again, by for example increasing it in magnitude when paired with the compounded CSs, blocking is not found. Thus, conditioning proceeds only when the CS provides unique, non-redundant, information about the US.

A number of theories similar to the Rescorla-Wagner model exist. Kamin (1969) earlier proposed that learning will only proceed when the US is surprising. Although Kamin's characterization is similar to the Rescorla-Wagner model, it differs in that the direction of the surprise (whether the animal was expecting more and got less or was expecting less and got more) is irrelevant in Kamin's model while it is important in the Rescorla-Wagner model. In Pearce and Hall's (1980) theory, the associability of a stimulus is high when that stimulus is followed by an unexpected US and low when followed by an expected US. All of these competing models use the same error-correction logic, whereby an estimate of the likelihood of the US given the CS is updated upon presentation of the CS and an unexpected outcome. The estimated likelihood of the US given the CS increases if an unexpected US follows, or decreases if an expected US fails to follow (Pearce & Bouton, 2001). The prevalence of this feature within models of Pavlovian conditioning indicates that, in some form, the degree to which a US was expected or unexpected prior to the CS is an important aspect of Pavlovian conditioning.

1.6 Conclusions

Generally, a biologically neutral stimulus will become able to maintain responding if it: first, signals something about forthcoming biologically relevant stimuli and second, does so uniquely and non-redundantly. Although the stimulus may have some acquired appetitive or aversive properties, these properties are not responsible for the behavioural effects of the stimuli.

Chapter II

Although a stimulus may have hedonic or appetitive properties, any increase in responding that accompanies or follows that stimulus may not be due to those appetitive properties. Evidence was presented in Chapter I that the behaviour engendered by conditional reinforcers may be better characterized as due to the discriminative rather than the hedonic properties of the stimuli. When a stimulus previously paired with a primary reinforcer is presented during operant responding for that primary reinforcer, responding increases to a greater degree than when the stimulus is presented during operant responding for a different primary reinforcer (Crombag, Galarce, & Holland, 2008; Galarce, Crombag, & Holland, 2007), suggesting that such stimuli contain "sensory-specific reinforcement information". Davison and Baum (2006; 2010) arrived at a similar conclusion and noted an intriguing implication: if the behavioural effects of conditional reinforcers are due to their discriminative rather than their motivational properties, perhaps the same is also true for the behavioural effects of primary reinforcers. Similarly, Shahan (2010) concluded his review of conditional reinforcement by noting that primary reinforcer effects may be also due to a signalling rather than a response-strengthening function.

Primary reinforcers are stimuli which are appetitive because of the animal's phylogenetic, rather than its ontogenetic, history. The discriminative functions of reinforcers have been recognized for years. Ferster and Skinner (1957) noted the familiar "scallop-pattern" seen in the cumulative response record when reinforcers are delivered according to a fixed-interval (FI) schedule. They attributed this to the signalling function of each reinforcer: a reinforcer in this context signals a temporary and predictable period of extinction. Reinforcers have a number of effects attributable to their signalling properties.

2.1 Choice behaviour

Herrnstein (1970) noted that all behaviour is choice behaviour: at all times organisms have multiple response options available to them. Moreover, the time that an organism allocates to a particular response is directly proportional to the relative reinforcers obtained for that response (generalized matching; Davison & McCarthy, 1988; Herrnstein, 1961). Such matching is commonly obtained in choice procedures and in fact itself suggests discriminative control by the relative reinforcer rates: the relative frequency of reinforcement *must*, in some sense, be discriminated in order for the relative response rates to match these relative reinforcer rates (Gallistel, 1990).

According to the contingency-discriminability approach to choice behaviour, matching comes about as a result of attempts to allocate behaviour in the same proportion as the perceived reinforcer ratio (Davison & Jenkins, 1985; Davison & Nevin, 1999). Deviations from perfect matching (commonly obtained; Baum, 1974) arise partly because the response-reinforcer relations are imperfectly discriminated: a reinforcer delivered contingent on a peck to Alternative 1 may be perceived as having come from Alternative 2. As the contingency is made more discriminable, by for example making the response alternatives more different, the response ratio becomes more similar to the reinforcer ratio (Alsop & Davison, 1991; Godfrey & Davison, 1998; Miller, Saunders, & Bourland, 1980). This suggests that matching of the response ratio to the reinforcer ratio requires discrimination of the relative reinforcer rates. Further, animals can report the responses that preceded the most recent reinforcer (Alsop & Davison, 1992; Jones & Davison, 1998; Killeen, 1978; Killeen &

Smith, 1984; Lattal, 1979). Thus, the contingencies of reinforcer delivery are indeed a discriminable feature of the environment.

2.1.1 Choice behaviour in transition

Moreover, the rate at which the reinforcer ratio changes appears to also be discriminable. The time that it takes for the behaviour ratio to adjust to a new reinforcer ratio is directly related to the frequency with which the reinforcer ratio changes. When the reinforcer ratio changes every 6 sessions, nonzero sensitivity to the previous reinforcer ratio is detectable 3 sessions after a change but not 6 sessions after (Davison & Hunter, 1979). When the reinforcer ratio changes every single session, the effect of a particular session's reinforcer ratio dissipates within 3 or 4 sessions (Hunter & Davison, 1985; Schofield & Davison, 1997). When the reinforcer ratio changes up to 7 times each session, sensitivity to the current reinforcer ratio reaches asymptote within 12 reinforcer deliveries (Aparicio & Baum, 2006; Aparicio & Baum, 2009; Davison & Baum, 2000, 2002, 2003; Krägeloh & Davison, 2003; Landon & Davison, 2001; Landon et al., 2003a). In a variety of procedures, behaviour generally appears to very quickly adjust to new contingencies when the reinforcer ratio often or frequently changes (Baily & Mazur, 1990; Gallistel et al., 2007; Gallistel, Mark, King, & Latham, 2001; Lau & Glimcher, 2005; Mark & Gallistel, 1994; Mazur, 1992, 1997; Palya & Allan, 2003).

Mazur (1997) provided direct evidence that the rate at which choice behaviour adjusts to new contingencies depends on the rate at which those contingencies change. He changed the reinforcer ratio after either 2 or 8 sessions and reported that behaviour adapted more quickly when the previous reinforcer ratio had been in place for a shorter period. This finding has also been obtained in a field study with free-ranging animals. Devonport and Devonport (1994) set up two feeder sites for chipmunks and ground squirrels. Only one of these feeders was baited at a time, and Devonport and Devonport varied the frequency with which the baited feeder alternated. The feeders were then removed for a period and later reintroduced. When the baited feeder infrequently alternated, the animals were more likely to revisit the last baited site. When the baited feeder more frequently alternated, the animals did not return to the most recently baited food site but instead visited the overall richer site. Devonport, Hill, Wilson and Ogden (1997) replicated this finding with rats in a radial arm maze: animals were more likely to revisit the most recently baited site when the food sites had previously been stable.

This dependence of the rate of behaviour change on the rate of environmental change is captured in a number of models which aim to predict behaviour in foraging situations (e.g., Cowie, 1977; Kacelnik, Krebs, & Ens, 1987). In such models, either the size of the memory window, or the relative weight given to the most recent event (versus a summary of older events), varies as a function of the rate of environmental variability. Thus reinforcers signal not only their own schedules of delivery, but also the rate at which those schedules change.

2.2 Control by local and global contingencies

The often-obtained findings that relative response rates quickly approach relative reinforcer rates has led some to propose that matching is inherent (Gallistel et al., 2007). Others, however, have argued against the primacy of matching, claiming that it is an artefact arising from the aggregation of many responses over time. According to *momentary maximizing* (Shimp, 1966, 1976a, 1976b), global-level matching results from a process operating at a local level whereby every response is allocated to the alternative with the momentarily higher probability of reinforcement. Shimp (1966) found both global matching in a concurrent schedule procedure, and consistent sequential dependencies in which each response was made to the alternative more likely to arrange the next reinforcer (see also Hinson & Staddon, 1983; Silberberg, Hamilton, Ziriax, & Casey, 1978). He concluded that global matching was a function of these sequential dependencies operating at a local level, and theorised that there is no direct control by the long-term distribution of reinforcers.

Conversely, Nevin (1969) obtained global-level matching in the absence of consistent molecular response patterns, and even some suggestion of responses emitted to the alternative locally less likely to produce the next reinforcer (although see Silberberg et al., 1978 for an argument that Nevin's analysis was insensitive to the detection of sequential dependencies). Since the same global patterns of behavioural allocation (matching) can apparently arise from a variety of different molecular response patterns, global matching may not in fact be derivative from a more fundamental process operating at a local level (Herrnstein & Loveland, 1975). Buckner, Green and Myerson (1993) also reported that while reinforcers did have consistent short-term effects, these short-term effects were not likely responsible for matching at the long-term, molar level. While a reinforcer increased the length of the visit during which it was delivered, there was little consistent effect on subsequent visits (c.f., Baum & Davison, 2004; Davison & Baum, 2003). Buckner et al. also found that time-allocation in 15-minute reinforcer-free periods, as well as in more extended periods of extinction, was indistinguishable from time-allocation outside of these extinction periods. Although reinforcers did have demonstrable local effects, these local effects were not essential to the long-term effects (matching). This again implies that matching is fundamental and not dependent on local-level regularities.

In a standard steady-state concurrent schedule procedure, Landon, Davison and Elliffe (2002) found that individual reinforcers generate a brief period of preference to the alternative which had provided that reinforcer, and that sequences of same-alternative reinforcers shifted preference towards the alternative providing those reinforcers. In addition to these local effects, Landon et al. also found control by longer-term aggregations of reinforcers: behaviour after the initial preference pulse stabilized at a level reflecting the overall reinforcer ratio, and the tree structures were shifted towards the alternative providing more reinforcers overall. All three of these findings (Buckner et al., 1993; Landon et al., 2002; Nevin, 1969) suggest order at both local and relatively+ extended levels.

Direct control by longer-term aggregations of consequences is also suggested by research demonstrating that animals can behave sub-optimally in the short-term if such behaviour is optimal in the long-term. Wanchisen, Tatham and Hineline (1988) and Neuman, Ahearn and Hineline (2000) both presented pigeons with a two-alternative independently-arranged concurrent schedule: one of the alternatives was a standard FR schedule and the other was a progressive ratio (PR) schedule in which the response requirement increased after each reinforcer. When the PR continually increased throughout the session, the pigeons responded on that alternative until the PR schedule requirement reached or slightly exceeded that of the FR. At this point preference switched to the FR for the remainder of the session. In another condition, completing one FR reset the PR. When this was the case, the pigeons responded to the FR even when its response requirement was still greater than that active in the current PR schedule. In other words, they responded to the alternative which momentarily provided a more delayed reinforcer in order to ensure a higher overall reinforcer rate, suggesting control by long-term consequences.

Control by relatively extended-level contingencies is also implicated in an experiment reported by Jones and Davison (1997). They arranged a concurrent schedule procedure with an additional contingency whereby if relative responding to the left alternative in the first portion of the session (the instrumental phase) exceeded .75, a different reinforcer rate would be operative in the second portion of the session (the contingent phase). As the rate of reinforcement contingent on exceeding response proportions of .75 increased, so too did the relative response rate in the instrumental phase. Thus, choice does appear to be sensitive to temporally extended events and this sensitivity need not be mediated by processes operating at a more local level.

This control by temporally extended events does not necessarily or logically preclude control by local-level contingencies. Choice procedures can be arranged in which control by local contingencies calls for one allocation of behaviour, whereas control by longer-term contingencies calls for a completely different allocation. Williams (1991a) reviewed a number of these such studies (Hiraoka, 1984; Zeiler, 1987) and concluded that preference in these procedures appears to lie somewhere in between the two extremes predicted by exclusive control at either a local or a global level. Additionally, when discriminative control by the local contingencies is degraded in such preparations, control by the longer-term contingencies emerges (Silberberg & Williams, 1974). Williams found that when discrimination of the local-level contingencies was high, control by the long-term contingencies was attenuated (extended sensitivity to reinforcement was about 0.56). However when control by the local level contingencies was degraded (by increasing the interval between reinforcers), control by the longer-term contingencies improved (sensitivity to reinforcement increased to 0.8-1.0). This suggests that independent control by local and global contingencies is possible and that apparent control by the latter is not dependent on control by the former.

2.3 Local preference when the local probability of a reinforcer is low

Although preference does appear to carry over from one interval into the next, even when there are clear discriminative stimuli indicating that the contingencies have changed (Davison & Baum, 2002), this tendency to behave as if the immediate future will be similar to the immediate past can be overcome. Krägeloh, Davison and Elliffe (2005) arranged a two-alternative concurrent schedule procedure and varied the probability of a same-alternative reinforcer (a continuation) across conditions. As the probability of a continuation (and hence the average length of a sequence of same-alternative reinforcers) increased, the probability of emitting the first peck after a reinforcer to the just-productive alternative also increased. Switching immediately to the not-just productive alternative was always more likely than staying when the probability of a same-alternative reinforcer was less than .35. The preference pulses to the just-reinforced alternative were also more extreme and lasted longer when the conditional probability of same-alternative reinforcer was higher. This was not only generally true, but was also true for preference after discontinuation reinforcers in particular. This result is noteworthy because the length of the preceding series of reinforcers on the not-just-productive alternative was, on average, longer when the probability of a continuation was higher. In more typical concurrent-schedule procedures, residual effects of the earlier reinforcers are present: preference to the just-productive alternative is typically attenuated when this reinforcer is preceded by a longer sequence of reinforcers from the other alternative.

Krägeloh et al.'s (2005) results thus confirm that animals can behave as if the immediate future will be unlike the immediate past: just because one alternative delivered the greater portion of the immediately prior reinforcers does not indicate that this alternative will continue to provide more reinforcers. Preference in this procedure was a function of the signalled probability of a same-alternative reinforcer, not of the preceding sequence of reinforcers. Davison and Baum (2006; 2010) similarly found that preference after a non-food stimulus (a magazine-light or a keylight illumination) was a function of the signalled probability of a same-alternative food. When this probability was high, preference was to the alternative that provided that stimulus. When however this probability was low, preference was towards the not-just-productive alternative.

2.3.1 Strict alternation procedures

Although they reported clear control by the local probability of a continuation, Krägeloh et al. (2005) also found some evidence of an invariant tendency to stay at the just-productive alternative. Even when the probability of a continuation reinforcer was 0 (indicating strict alternation of reinforcers across the alternatives), the probability of staying at the just-productive alternative immediately after a reinforcer remained greater than zero (Krägeloh et al.'s Figure 5). In fact, for 4 of the 5 pigeons, the probability of staying on the right after a right reinforcer exceeded .5 even when there was no chance of a right reinforcer in this period. Thus, although there was some control by the probability of a continuation, this control was by no means perfect.

Strict alternation appears to be regularly obtained with difficulty. Shimp (1976b) arranged win-stay lose-shift and win-shift lose-stay contingencies. In a win-stay, lose-shift preparation, a reinforcer signals that the next reinforcer is highly likely to appear on the same alternative (win-stay). If a trial does not end in reinforcement, the next reinforcer is highly likely to be on the other alternative (lose-shift). These contingencies are reversed in win-shift, lose-stay preparations. Although performance in both tasks was good, adherence to the win-stay contingencies was always greater than adherence to the win-shift contingencies at all inter-trial intervals. The proportion of correct win-stay responses was .96, .84 and .74 for delays of 2.5, 4.0, and 6.0 s respectively, compared to .87, .81 and .55 for win-shift responses.

In an early experiment, Hearst (1962) found good evidence of strict alternation in a discrete-trials procedure even at relatively long (10-s) delays. Accuracy was always above 90% for 3 of his 8 pigeons, was consistently greater than 75% for 6 pigeons, and was always greater than 60% for all pigeons. However no control conditions or comparisons were conducted to compare strict alternation performance with performance in any other procedure. Williams (1971a; 1971b) similarly arranged strict alternation in a discrete-trials procedure but classified the first key to receive either 15 or 30 pecks (varied across subjects) as the "responded to" key in that trial. If an animal initially emitted fewer than 15 or 30 responses on the just-productive alternative, and then switched to the other alternative (where it completed the response requirement) that trial was coded as "correct". Although percent correct decreased as the delay interval increased, performance was always above 55% correct for all 4 pigeons even at the longest delay (45 s).

Greater-than-chance performance in Williams' (1971a; 1971b) task depended on the FR requirement: it was achieved when FR 15 or 30 were used but not when FR

1 or 5 were used. Williams noted that this might have been because of some tendency to repeat the just-productive response immediately after food. Small FR values often resulted in the FR requirement being completed and that trial being classed as "incorrect". However when a larger FR requirement was in place, the first pecks to the just-productive alternative did not complete the FR. There may be some initial tendency to stay at the just-productive alternative before responding can switch to the other alternative, where reinforcement is arranged.

The consistent pattern of first responding to the just-productive alternative before switching to the other alternative suggests that the strict alternation contingencies in Williams' (1971a; 1971b) experiments may not have acquired any control: the pigeons may have also initially responded to the just-productive alternative before switching to the not-just-productive alternative if reinforcers randomly, rather than strictly, alternated. Williams did not report the results of any such control condition, leaving open the possibility that the overall distribution of reinforcers was responsible for the distribution of responses in this procedure. As noted above, behaviour can be directly controlled by temporally extended distributions of reinforcers. Comparison with control conditions arranging the same overall 1:1 reinforcer ratio and no strict alternation would better allow for the attribution of particular response patterns to discrimination of the strict alternation of reinforcers. Additionally, any oscillations in preference through time (initially preferring the just-productive alternative before turning to preference for the other alternative) would more easily be seen in a free-operant, rather than a discrete-trials, preparation. Krägeloh et al.'s (2005) experiment was conducted in such a free-operant situation, and contained a number of conditions with different probabilities of a continuation reinforcer. As in Williams' (1971a; 1971b) studies, Krägeloh et al.

also reported that the probability of staying immediately after a reinforcer was high, even when the probability that the next reinforcer would be on that alternative was 0. Also just as Williams reported, Krägeloh et al. found that preference turned to the not-just-productive alternative after the initial responses to the just-productive alternative.

Thus, although animals can learn to respond to the not-just-productive alternative, control by such contingencies appears to be imperfect. Davison, Marr and Elliffe (in press) varied the likely location of the next reinforcer as a function of behaviour prior to that reinforcer. In two parts of their experiment, they arranged negative feedback functions such that the behaviour ratio in one inter-reinforcer interval (IRI) determined the relative reinforcer probability in the next IRI: If the behaviour ratio was extreme towards one alternative in an IRI, the next reinforcer was more likely to be on the other alternative. When the slope of the feedback function was -3, the probability of a continuation was lower than when this slope was -1.

Davison et al. (in press) conducted multiple linear regressions of relative choice in each IRI as a function of relative choice in the last IRI, and the location of the last reinforcer. They found relatively weak control of the behaviour ratio in an IRI by the behaviour ratio in the previous IRI. This is not to say that there was no effect: there was a clear difference between Phase 1, which arranged a typical two-alternative concurrent schedule procedure (and the associated positive feedback function) and the phases with negative feedback functions. Also, increasing the slope of the negative feedback function (from Phase 2 to Phase 3) significantly decreased the effect of the prior reinforcer. These results again suggest that, although something approaching strict alternation can be obtained given the appropriate contingencies, such preference reversals across IRIs (which is required in all strict alternation procedures) are obtained with difficultly.

The previously reported problems with attaining perfect strict alternation may thus have been due to carryover of preference from one IRI into the next. Perfect adherence to strict alternation procedures requires maximal control by the current contingencies and minimal preference carryover from the previous IRI. Minimizing such carryover appears to be difficult and may in fact be impossible. Preference in the frequently-changing procedure regularly carries over from one component into the next despite inter-component blackouts signalling the end of the component (Davison & Baum, 2002; Landon & Davison, 2001). Additionally, preference often carries over from one session to the next despite extended experience with the independence of daily reinforcer ratios (Hunter & Davison, 1985; Schofield & Davison, 1997).

The problem does not appear to be a result of previous behaviour acquiring discriminative control over current behaviour: as discussed above this appears to be readily accomplished (Alsop & Davison, 1992; Jones & Davison, 1998; Killeen, 1978; Killeen & Smith, 1984; Lattal, 1979). The problem instead appears to be one of countering some tendency to repeat the previous response. Shimp (1966) calculated the probability of making a response on each trial given the probability of making that choice on previous trials. In all cases, the probability of making a choice on one trial was greater if that choice had been made on the previous trial, regardless of whether or not that prior choice had been reinforced. Schneider and Davison (2005) also reported some bias to making repeat responses (left-left or right-right) when two-response sequences (left-left, left-right, right-left and right-right) were reinforced in a 4-alternative concurrent schedule procedure. This tendency towards perseveration, or for preference to carry over from one interval to the next makes assessing the discriminative functions of reinforcers difficult.

2.4. Alternation of sequences and reinforcer counting

While single alternation is apparently obtained or at least approximated (at times with difficulty) in a number of procedures with nonhuman animals, alternation of sequences (e.g., three left reinforced responses followed by three right reinforced responses) seems much more difficult to obtain. In both single- and sequential-alternation procedures, recall of the response that preceded the last reinforcer is required. Sequential alternation procedures however also require recall of the response that preceded earlier reinforcer(s). Williams (1976) found no evidence of learning a double alternation procedure after 20 sessions. Rather, 5 of the 6 pigeons perseverated on an alternative until they did not get a reinforcer, and then switched. Evidence of adherence to the double alternation procedure was only obtained after 135 sessions. Even then, accuracy was a decreasing function of the number of preceding correct trials. For some animals, percent correct dropped from greater than chance performance (75% correct) to near-zero levels in one step.

In a review of the literature on double alternation, Kundey and Rowan (2009) concluded that, while humans can perform such tasks with ease (adult humans being better than children), even nonhuman primates demonstrate considerable difficultly with these tasks, requiring longer to reach criterion. Non-primates have even more trouble acquiring the task, completely failing in most investigations. In the few cases where they have acquired it, they failed to extend beyond the trained sequences, suggesting that the animals were using complex behaviour chains rather than any strict alternation rule. For example in Williams' (1976) double alternation procedure, percent correct was high if there had been 0 to 3 preceding correct trials but fell drastically after 4 preceding correct trials. This suggests that the pigeons were emitting a complex behaviour chain (e.g., left response, left response, right response, right response) rather than responding based on the number of preceding same-alternative reinforcers. Humans working on such tasks verbalise their behaviour as sequential alternation (Gellermann, 1931), suggesting a conceptual understanding based on the preceding number of same-alternative responses (e.g., two left followed by two right) rather than simple memorisation of long sequences.

In their experiment, Kundey and Rowan (2009) required rats to nosepoke in either single or double alternation. At the start of a trial, a light in each receptacle was lit. Upon an error, the incorrect receptacle was darkened. This procedure proved successful at producing double alternation. Kundey and Rowan looked separately at the errors on Response 1 and Response 2 (for the single alternation group) and on Switch to 1, Repeat 1, Switch to 2 and Repeat 2 for the double alternation group. The probability of making any error was low for the single alternation group (.06). For the double alternation group, the probability of making a repeat error (not reported but less than .05 according to their Figure 3) was noticeably lower than the probability of making a switch error (.183). Although this indicates some tendency to stay, the number of switch errors was still significantly fewer than predicted by chance, indicating that the double alternation task was acquired to some degree. Kundey and Rowan's Experiment 2 confirmed the importance of the correction procedure: a group with no such procedure failed to acquire the task. Thus, with rats at least, some external stimulus apparently needs to mark incorrect responses in order for correct responding to emerge. The role of the overall probability of a continuation, identified by Krägeloh et al. (2005) as a controlling variable, was not addressed in this study.

It thus appears difficult to train an animal to use the preceding number of same-alternative reinforcers as a discriminative stimulus, if this number of preceding same-alternative reinforcers is greater than 1. The literature on numerical competence in nonhuman animals suggests that although nonprimates may not "count" according to a formal definition (Davis & Perusse, 1988), they do have rather sophisticated numerical abilities (Dehaene, 1997). For example, Capaldi and Miller (1988a) reported that rats could both keep track of how many total foods they had eaten and, at the same time, keep track of how many of each of two qualitatively different foods they had eaten. Across a series of experiments, rats were required to run through a series of runways, some of which ended in food (R), others of which did not (N). Sequences of NRRN and RRN were interspersed randomly with the time held in each runway varied (to ensure that neither ITI nor time since the first food signalled the end of a sequence). Rats always ran more slowly after the second food, indicating that they were able to anticipate that the final runway would not contain food. An experiment in which the sequences were extended to NRRRN and RRRN confirmed that the animals were keeping track of the number of consumed foods. Animals appear to make discriminations based on the numerical features of the situation even when not required to do so and even when other features of the situation can be equally used (e.g., animals use numerical as well as temporal features when both are relevant; Fetterman, Stubbs, & Dreyfus, 1986; Roberts & Mitchell, 1994). This suggests that discriminations based on the number of items are indeed possible and may even be highly likely.

2.5 Foraging and counting foods

Dehaene (1997) noted the evolutionary advantages of numerical competence. For example, an animal capable of discriminating a food site with 4 foods from a food site with 3 will be able to preferentially exploit the richer food site. More directly related to the above research, an animal able to keep track of how many prey have been consumed within a patch containing a fixed number of prey will be able to leave that patch as soon as it is empty. Such a numerically competent animal will not abandon a patch that still contains some prey, nor will it remain in a patch after consuming all foods within. Gibb (1966) hypothesized that animals foraged in just such a manner. He collected the pine cones from which titmice had harvested moth larvae. The titmice create holes in the pine cone to access their prey. These holes can be counted as a measure of the number of consumed prey per patch. The remaining larvae can also be counted, allowing for a measure of the proportion of larvae consumed per pine cone. Occasionally, a tree was more heavily infested relative to the surrounding trees. The birds took proportionally fewer larvae from each pine cone in these more heavily infested trees. This suggests that the titmice learned to expect a particular number of larvae in each pine cone and, once that number was obtained, abandoned the pine cone.

Krebs, Ryan and Charnov (1974) noted that this interpretation is problematic. First, Gibb's (1966) conclusions were based on observation of the artefacts of behaviour (the holes in the pine-cones), not of the actual behaviour (foraging). Second, hunting by expectation could only develop as a strategy if each pine cone contained a standard number of larvae. Variability in the number of prey per patch would inhibit the formation of any expectations, and Gibb himself provided the evidence of variability in prey density. Krebs et al. arranged for black-capped chickadees to forage for mealworms hidden in artificial pine cones. When the birds encountered a long sequence of patches, each containing the same number of prey per patch, they did not learn to expect that fixed number of prey in each patch (Gibb's hypothesis of hunting by expectation); they also did not learn to spend a constant time in each patch, although the data on this were less clear. However the birds did have a constant giving-up time for all patches within an environment: the interval from the time of the last prey capture until the patch was abandoned was constant and inversely related to the average food rate in that environment. Rather than counting the number of foods consumed in order to decide when to leave a patch, these animals apparently decided when to leave by timing the interval since the last prey capture. This lack of control by the number of prey per patch was replicated by Roche, Timberlake, Glanz and Stubbs (1998).

Another noteworthy example of failing to obtain control by the number of prey per patch, despite this number being constant, was obtained by Lima (1984). Lima's results are particularly intriguing because, although they suggest a lack of control by the number of consumed foods per patch, they do suggest some control by the number of prey per patch. He arranged a some-patches-are-empty paradigm where downy woodpeckers were presented with planks of wood with 24 holes into which (cryptic) food could be placed. Some patches (planks), contained no food in any hole. Across conditions, Lima varied the (fixed) proportion of holes filled with food in the non-empty patches. He reported that the number of holes sampled in an empty patch was inversely related to the proportion of holes containing food in non-empty patches: more holes were sampled in empty patches when non-empty patches were a quarter-filled than when they were half-filled. Roberts (1991) replicated these results with rats in a modified radial arm maze. There were 4 arms with 10 holes per arm that could contain food. In each session, 2 of these 4 arms (patches) were empty. Across conditions, Roberts varied the number of holes containing food in the non-empty arms (1, 3, 5, 7, 10) and found the same result as Lima: the number of feeders visited in empty arms was clearly related to the number of foods in the nonempty arms.

These results suggest some control by the number of prey per patch. However behaviour in non-empty patches suggested no such control: the birds sampled all holes in the baited patches even after all prey had been found. Whether the baited patch was completely full, half-full or a quarter-full, the modal number of holes opened per non-empty patch was 24 (all holes). Even when only two holes per nonempty patch had food (a small, easy to discriminate number), the animals continued to sample after consuming all foods. Lima (1984) also conducted an informative probe: during one session in the half-full condition, all foods in the non-empty patches were concentrated in the bottom half of the patch. The animals were generally systematic (sampled sequentially). Even when all of the foods were clumped, the animals continued to open most of the holes. Thus, although the animals did have some estimate of patch quality, as evidenced by an inverse relationship between the number of filled holes in non-empty patches and the number of holes visited in empty patches, this estimate was neither based on counting nor on the number of consumed prey in the current patch: once a patch was found to contain food, the animal continued to exploit it until all holes were open, even after all foods were captured and even when there were only 2 foods per patch. Without counting, how did they track the number of holes to visit before abandoning a patch as empty? Kamil and Roitblat (1985) suggested that the birds had some estimate of average prey density but could not keep track of the individual counts within a patch. This suggests some control by the global contingencies (the average prey density per patch) and no control by more local contingencies (the local prey density in the current patch as signalled by the number of foods already obtained in that patch).

2.6 Bayesian foraging

A foraging animal's behaviour as controlled by both long-term and more immediate consequences can be described as a Bayesian process. Bayes' theorem explains how an estimated prior probability of an event is transformed by new information into an estimated posterior probability (McNamara, Green, & Olsson, 2006; Olsson & Brown, 2006; Olsson & Holmgren, 1998; Stephens & Krebs, 1986). According to this view, animals arrive in a foraging situation equipped with prior estimates of the resource. One potential source of these prior estimates could be previously-experienced food distributions. Additional information acquired whilst sampling is then incorporated with the prior information to arrive at a posterior estimate of patch quality. Behaviour at these times may be understood as a function of more local contingencies. Bayesian models predict how animals combine previous experience with sampling information.

If, for example, patches within a habitat are known to contain either 0 or 2 prey items with equal probability (as in some conditions of Lima's, 1984 experiment), the prior probability of the patch containing food is .5. After the animal encounters a prey item however, the posterior probability that the patch contains prey becomes 1.0. Other, non-food, sources of information can also update probability estimates. In this example, as time in the patch without finding any prey increases, so too does the probability that the patch is empty. Other stimuli can also serve as information about the probability of a patch containing prey**.** Stephens and Krebs (1986) give the example of a predator hunting for animal prey. Rustling in the grass updates the probabilities and signals that the patch is more likely to contain prey.

Bayesian updating is often touted as uniquely predicting that experience prior to the current patch influences behaviour. Valone (2006) reviewed the literature and concluded that animals generally, though not always, behave in ways consistent with Bayesian foraging models: animals do appear to base their decisions not just on current sampling information, but also on the prior probabilities. Incorporating information about the prior distribution of events leads to more accurate estimates than would be obtained by simply relying on local information. Put another way, the long-term or global distribution of the resource, in addition to its local distribution as signalled by some event, influences behaviour. Lima's (1984) study indicates some control by the prior probability of a reinforcer: the number of holes sampled before abandoning a patch as empty increased as the number of holes per filled patch decreased. Valone (1992) also reported some influence of prior distributions: each patch contained 12 "flowers" which could contain food for the black-chinned hummingbirds. Either 2 or 5 of the flowers per patch contained food and Valone (1992) found that the number of probes prior to departure decreased as the number of consumed prey increased. The inverse relationship between the number of already-consumed foods in this patch and the probability of finding a further food was discriminated. Put another way, there was control by the prior reinforcer density.

Bayesian perspectives on foraging thus lead to many of the same conclusions that follow from accounts which consider behaviour a joint function of long-term and more local contingencies. In the absence of a discriminative stimulus signalling that local contingencies are in effect, behaviour will be a function of longer-term contingencies. In the terminology of Bayesian foraging, the prior probability estimates will govern behaviour until some informative event changes these probability estimates.

The reliability of a stimulus is the probability that it correctly identifies the environment as being in a particular state (Koops, 2004). Reliability should enhance the value of a signal — the more reliable the information, the more valuable the signal. If, for example the probability of the patch rustling is as high when there is a prey item in that patch as when it is empty, the signal's reliability will be low. A signal's value also depends on the level of certainty in the environment. If, for example, the environment is equally likely to be in one of two states, a signal indicating the current state is more valuable than it would be if the environment is in one of the states 99% of the time. In the latter environment, an animal could behave appropriately most of the time while ignoring the stimulus. Such inattention to the stimulus would lead to inappropriate behaviour about half of the time in the equal-probability environment. In a discrete-trials concurrent schedule procedure, McLinn and Stephens (2006) varied the overall reinforcer ratio and, on each trial signalled the alternative more likely to provide the next reinforcer. The Blue Jays were more likely to respond to the alternative signalled more likely to provide the next reinforcer when the overall probability of a reinforcer on that alternative was .5. When one alternative was overall more likely to provide the next reinforcer ($p = .75$), this alternative was always preferred, even if a signal indicated that the other alternative was more likely to provide the reinforcer on this trial. However this result could also have been due to excess control by the global contingencies of reinforcement overshadowing any control by the local contingencies.

2.7 Summary and conclusions

The control that reinforcers and their circumstances of delivery have on behaviour can be seen as discriminative. The common finding that global relative response rates approximately match global relative reinforcer rates (Davison & McCarthy, 1988) may result from such discriminative control by reinforcers (Davison

& Jenkins, 1985; Davison & Nevin, 1999). Additionally, control by local contingencies, such as in strict alternation procedures, may also result from the discriminative properties of reinforcers. Behaviour generally more closely approximates the contingencies of reinforcement when these contingencies are more clearly discriminated. Apparent failures of behaviour to closely approximate the contingencies of reinforcement may arise from shortcomings in discrimination of the local or global contingencies. This may account for the apparent difficulty in sequential alternation and other procedures that require discrimination of the number of consumed reinforcers from one location. Aside from these issues of impaired discriminative control, there also appears to be some general reluctance to switch, or some general tendency for preference to carryover from one epoch to next. These issues notwithstanding, conceptualizing of reinforcer control as discriminative allows for a coherent and complete characterization of a number of disparate findings.

Chapter III

3.1 Experiment 1

Davison and Baum (2006) found that arranging a positive correlation between the left: right reinforcer ratio and the left: right non-food stimulus ratio led to a local increase in preference to the alternative that produced that non-food stimulus. When this correlation was negative, preference in the post-stimulus period was for the other, not-just-productive alternative. Contrary to a pairing hypothesis account of conditional reinforcement (Williams, 1991b, 1994), whether the stimulus had never been paired with food (green keylight) or had an extensive history of pairing with food (magazine-light) made no difference to the local effects of the stimulus. The crucial factor in determining the direction of local preference in the post-stimulus period was the correlation between the stimulus and food ratios.

Davison and Baum (2006) conducted their study in a frequently-changing environment where the reinforcer ratio changed up to 7 times per session. In this procedure, there is a degree of uncertainty about the current reinforcer ratio: sensitivity to the current component reinforcer ratio (rate of change in the log behaviour ratio that occurs as a function of changes in the log food ratio; Baum, 1974, 1979; Lobb & Davison, 1975) typically starts at or below 0 and asymptotes at about .4-.7 after about 8 reinforcer deliveries (Davison & Baum, 2000; Landon & Davison, 2001). In steady-state procedures on the other hand, data analysis does not typically commence until either a specified stability criterion has been met, or there has been sufficient time in the current condition for it to be reasonably assumed that the behaviour ratio has stabilized (Landon et al., 2002). Unlike in frequently-changing procedures, the current reinforcer ratio is likely to be highly discriminated throughout the data-collection period in these steady-state procedures.

The local effects of reinforcers first reported in frequently-changing environments (Davison & Baum, 2000, 2003; Landon & Davison, 2001) have been replicated in steady state preparations (Landon et al., 2002; Landon, Davison, & Elliffe, 2003b). *Continuations,* defined as uninterrupted sequences of reinforcers from one alternative, have diminished effects on preference as the number of same-alternative preceding foods increases, in both frequently changing and steady-state environments. *Discontinuations* (reinforcers obtained from the other alternative after a sequence of same-alternative reinforcers) have more extreme effects than do continuations in both frequently-changing and constant environments. Additionally, *preference pulses*, or short-lived increases in relative responding to the alternative that produced the last reinforcer, were first reported in a frequently-changing procedure (Davison & Baum, 2002) but have since also been reported in steady-state procedures (Landon et al., 2002, 2003b). Food reinforcers thus appear to have the same short-term effects regardless of the rate of environmental variation. Experiment 1 replicated Davison and Baum's (2006) procedure in the steady-state in order to determine whether conditional reinforcer effects are also unaffected by the rate of environmental variation.

If response-contingent stimulus presentations (correlated with food presentations) were the only source of information on the current food ratio in Davison and Baum's (2006) experiment, their finding would doubtlessly be replicated in the steady-state. However, the fact that sensitivity to reinforcement increases after each reinforcer in frequently-changing environments demonstrates that reinforcers also signal their own future contingencies. In Davison and Baum's (2006) experiment, these contingencies of food delivery changed up to seven times per session. In Experiment 1, these contingencies remained unchanged over many sessions. Thus, sensitivity to reinforcement will likely be asymptotically high by the time data collection begins in Experiment 1, leaving no room for successive foods or stimuli to increase measured sensitivity.

There were many similarities between the present experiment and Davison and Baum's (2006) experiment. First, as in Davison and Baum's experiment, response-contingent non-food stimuli were inserted into a concurrent schedule of food reinforcement. Second, food ratio-stimulus ratio correlations of +1, -1 and 0 were arranged and were varied across conditions. Third, stimuli both paired and unpaired with food were arranged. Experiment 1 differed from Davison and Baum's (2006) experiment in two major ways. First, in Experiment 1, the same left: right food ratio remained in place for at least 50 sessions, whereas the food ratio changed up to 7 times per session in Davison and Baum's experiment. Second, the stimuli used in the present experiment differed from those used by Davison and Baum. The unpaired stimuli were very similar across the two studies (red keylight illumination in Experiment 1 and green keylight illumination in Davison and Baum's experiment). The paired stimuli differed in a more noteworthy way. While Davison and Baum used magazine light as their paired response-contingent stimulus, red keylight illumination was used as the response-contingent stimulus in both the paired and unpaired conditions of Experiment 1. Because a pigeon would be unlikely to attend to the red keylight while consuming food, the stimulus in Experiment 1 (red keylight) was forward-paired with food, while the conditional reinforcer (magazine light) in Davison and Baum's experiment was simultaneously paired with food. In Experiment 1, all unpaired conditions were conducted before any of the paired conditions to ensure that the stimuli in the unpaired conditions had no history of being paired with food.

In Davison and Baum's (2006) experiment, when the food ratio-stimulus ratio correlation was nonzero, a response-contingent non-food stimulus was informative about the response likely to produce the next food: when the food ratio-stimulus ratio correlation was +1, food was likely to be on the just-productive alternative; when this correlation was -1, food was likely to be on the not just-productive alternative. The stimuli in Experiment 1 were equivalent to the stimuli in Davison and Baum's experiment in that the same stimulus ratio-food ratio correlations of -1, 0 or +1 were used. The different rates at which the reinforcer ratios changed across the two experiments meant that despite the same correlations being used in both, the stimuli in Davison and Baum's experiment were more likely to provide novel, nonredundant information about the current food ratio. As discrimination of the current food ratio is likely to be asymptotically high in Experiment 1, any information provided by the stimuli is more likely to be redundant. Unlike in Davison and Baum's experiment, there may be no need to attend to the correlation of the stimulus presentations with food, and hence no effect of whether this correlation is +1, -1, or 0.

3.2 Method

Subjects

Five adult homing pigeons, numbered 61, 62, 63, 64, and 66, all with extensive previous experience on two-key concurrent schedules (Krägeloh et al., 2005; Landon & Davison, 2001; Landon et al., 2003a) served in the present experiment. The pigeons were maintained at 85% ± 15 g of their free-feeding body weights by post-session supplementary feedings of mixed grain when required. Water and grit were freely available in the home cages at all times. The home cages were situated in a colony room with about 80 other pigeons.

Apparatus

Each pigeon's home cage also served as its experimental chamber. Each cage measured 375 mm high, 375 mm wide and 380 mm deep. Three of the walls were constructed of metal sheets while the fourth wall, ceiling, and floor were metal bars. 75 mm above the floor were two wooden perches, one positioned parallel to, and the other at a right angle to, the back wall. Three 20-mm diameter circular translucent response keys were positioned on the right wall. The keys were centred 114 mm apart and 223 mm above the perches. The centre key was never lit and remained inoperative throughout the study. When the two side keys were transilluminated white they required a force exceeding approximately 0.1 N to register an effective response. Also on the right wall, 138 mm below the centre key, was a 45-mm high by 45-mm wide food-magazine aperture. During food delivery, a hopper containing wheat, situated behind this aperture, was raised and illuminated for 3 s. All experimental events were arranged and recorded on an IBM-PC® compatible computer running MED-PC IV® software situated in an adjacent room.

Procedure

Initially, subjects were placed on a concurrent variable-interval (VI) variable-interval schedule of food reinforcement and, over the course of approximately one month, the overall VI schedule was progressively increased to VI 27 s, and a 2-s changeover delay (COD) was introduced. In Phase 1, the response-contingent stimuli (when present) were unpaired with food; in Phase 2 they were paired with food. Phase 1 also contained several conditions in which food was the only response-contingent event. Table 3.1 shows, for each condition, the overall schedule that produced food and stimulus events, the left:right food ratio, and the type of stimuli (paired or unpaired).

In all conditions in Phase 1 except Conditions 11 and 12, a single VI 27-s schedule arranged all response-contingent events (food and stimuli). When a response-contingent event was arranged, it was probabilistically determined ($p = .5$) whether that event would be food or a stimulus presentation, which consisted of a 3-s red illumination of the pecked key. (In conditions arranging no stimuli the probability of a stimulus was 0.) After a food or a stimulus was arranged, the contingencies appropriate to that condition determined whether that event would occur on the left or on the right key. For example, in 9:1, food-stimulus $r = +1$ conditions, a food would be arranged on the left key with a probability of .9 and a stimulus would also be arranged on the left key with a probability of .9. In 9:1, $r = -1$ conditions, a food would be arranged on the left key with a probability of .9, but a stimulus would be arranged on the left with a probability of .1. In this way, the alternative providing the greater number of stimuli could either be the alternative providing the greater number of foods or the alternative providing fewer foods. In all $r = 0$ conditions, the alternatives provided equal numbers of stimuli. Conditions 11 and 12 arranged an overall event-delivery schedule of VI 54 s and food was the only response-contingent event.

In Phase 2, the overall event-delivery schedule was changed to VI 24 s in order to keep the overall reinforcer rate the same as in Phase 1 (this change was done because of the 3-s stimulus presentations that preceded each food delivery in this phase). All events in Phase 2 were allocated in the same way as in Phase 1.

Table 3.1

Sequence of conditions in Experiment 1, along with the overall schedule of response-contingent events, the left: right food ratio, the type of stimuli (paired or unpaired) and the food-stimulus correlation in that condition in Phases 1 and 2.

Phase	Cond.	Overall Event Schedule	L:R Food Ratio	Stimuli
1	1	VI 27 s	1:1	None
1	2	VI 27 s	9:1	None
1	3	VI 27 s	9:1	Unpaired, $r = +1$
1	4	VI 27 s	9:1	Unpaired, $r = -1$
1	5	VI 27 s	1:9	Unpaired, $r = +1$
1	6	VI 27 s	1:9	Unpaired, $r = -1$
1	7	VI 27 s	9:1	Unpaired, $r = 0$
1	8	VI 27 s	1:9	Unpaired, $r = 0$
1	9	VI 27 s	1:9	None
1	10	VI 27 s	1:1	None
1	11	VI 54 s	1:9	None
1	12	VI 54 s	9:1	None
2	13	VI 24 s	1:1	Paired, $r = 0$
2	14	VI 24 s	9:1	Paired, $r = 0$
2	15	VI 24 s	1:9	Paired, $r = 0$
2	16	VI 24 s	9:1	Paired, $r = -1$
2	17	VI 24 s	1:9	Paired, $r = -1$
2	18	VI 24 s	9:1	Paired, $r = +1$
2	19	VI 24 s	1:9	Paired, $r = +1$

In Phase 1 (unpaired stimuli), when food was delivered, the left and right keylights were turned off, the food hopper was raised, and the magazine light was illuminated for 3 s. Response-contingent stimulus presentations consisted of the pecked key changing from white to red, and the other keylight turning off, for 3 s. After both reinforcement and keylight presentation, both keys were illuminated white and the contingent event schedule (which did not time during either response-contingent stimulus presentation or food presentation) started anew. Responses, if they occurred during a response-contingent stimulus presentation or during a food presentation, were not recorded in Phase 1.

In Phase 2 (paired stimuli), when food was delivered, first the productive-keylight was changed from white to red for 3 s, while the other key remained white. After this period, both keylights were turned off, the hopper was raised, and the magazine was illuminated for 3 s. During stimulus-alone presentations, the pecked key changed from white to red for 3 s and, unlike in Phase 1, the other key remained white. After either food or stimulus-alone presentation, both keylights were white and the contingent-event schedule proceeded to arrange events. Responses occurring during the stimulus or food delivery were recorded in Phase 2.

Sessions were conducted 7 days per week and all conditions lasted for at least 50 sessions. Data from the last 35 sessions of each condition were analyzed. The first 15 sessions of each condition were treated as transition data and were not analyzed. Towards the end of the experiment, overall response rates began to decrease substantially, and for this reason there were 9 sessions between Conditions 17 and 18 in which the stimuli were removed and the overall food rate was increased, bringing response rates to a level similar to those in the rest of the experiment. At this time, it was also decided, after observing the bird, to provide Pigeon 61 (which had by far the lowest response rates) with two non-contingent foods at the start of every session. These foods were not recorded and were intended purely to ensure responding within the session. Subsequently, this procedure was also implemented for Pigeons 64 and 62 at 9 and 22 days into Condition 19, respectively.

3.3. Results

3.3.1 Extended-level analyses

The total responses to each alternative in the last 35 sessions of each condition were summed and log (left/right) response ratios were calculated. The numbers of foods obtained from each alternative in the last 35 sessions of each condition were also recorded. From these data, sensitivity to reinforcement and bias (Baum, 1974) were calculated for each type of condition from the generalized matching law

$$\log \frac{B_1}{B_2} = a \log \frac{R_1}{R_2} + \log c \ , \qquad (3.1)$$

where *B* and *R* are responses and reinforcers on Alternatives 1 and 2 respectively. The parameter *a* is sensitivity to reinforcement, and measures the change in the response ratio that occurs as a function of a unit change in the obtained reinforcer ratio. Log *c* is bias, a constant proportional preference for an alternative. Single sensitivity and bias values were calculated using the 2, 3 or 4 conditions with the same food ratio-stimulus ratio correlation and pairing relationship with food, varying the left: right food ratio. These sensitivity to reinforcement and bias values are presented in Table 3.2. Of the 8 values presented for each subject, only 2 were obtained with more than 2 data points per estimate. Four conditions were conducted with no response-contingent stimuli and a VI 27-s schedule of food presentation, and 3 conditions were conducted with paired stimuli uncorrelated with food. Of the remaining 6 condition types, 2 conditions were conducted, one with a 9:1 food ratio, and the other with a 1:9 food ratio. Proportions of variance accounted, for the 2 condition types in which such a measure could sensibly be calculated, were between .89 and 1.0, with 6 of the 10 measurements between .99 and 1.0. Thus, sensitivity and bias can be considered well estimated.

Table 3.2

Individual-subject sensitivity to reinforcement and bias across condition types

Pigeon		No stimuli VI 27s	No stimuli VI 54s	Un-paired *r* = +1	Un-paired *r* = -1	Un-paired *r* = 0	Paired *r* = +1	Paired *r* = -1	Paired *r* = 0
61	*a*	1.07	1.20	1.08	0.91	0.95	1.12	0.58	0.86
	log *c*	-0.23	-0.09	-0.08	-0.15	-0.18	-0.08	-0.24	-0.22
62	*a*	0.86	1.02	1.00	0.88	0.86	0.84	0.65	0.60
	log *c*	-0.15	0.05	0.11	0.01	0.01	0.04	0.03	0.09
63	*a*	1.19	0.94	0.94	0.82	0.81	0.83	0.84	0.73
	log *c*	-0.35	-0.13	-0.08	0.00	-0.09	0.01	-0.05	-0.09
64	*a*	0.84	0.90	0.65	0.73	0.80	0.85	0.66	0.76
	log *c*	0.07	0.16	0.25	0.19	0.20	0.21	0.31	0.24
66	*a*	0.73	0.72	0.76	0.63	0.68	0.59	0.76	0.71
	log *c*	-0.14	-0.05	-0.07	0.02	-0.03	0.04	-0.12	-0.02
Mean	*a*	0.94	0.96	0.89	0.80	0.83	0.86	0.70	0.77
	log *c*	-0.16	-0.01	0.02	0.02	-0.01	0.03	-0.03	-0.02

In order to determine whether any of the manipulations had any detectable effects on sensitivity of the behaviour ratio to the food ratio, a repeated measures two-way analysis of variance was conducted with pairing condition and food-stimulus correlation as the factors. Neither main effect, nor the interaction of pairing and correlation, was significant ($p > .05$), indicating no effect of either pairing or correlation on the extended-level distribution of behaviour.

The number of response-contingent stimuli obtained from each alternative in each condition was also recorded. Using these data and the log response ratios, sensitivity of the log behaviour ratio to the left: right stimulus ratio was calculated. Separate sensitivity values were calculated for paired- and unpaired-stimulus conditions. Conditions with the same food ratio were used in the calculation of a single sensitivity value. Any sensitivity to the response-contingent stimulus ratio significantly different from 0 would indicate some control over changes in the behaviour ratio by the stimulus frequency ratio. Wilcoxon matched-pairs signed-ranks tests were conducted on the individual-subject data, comparing sensitivity to the contingent stimulus ratio in each condition type to 0. There were no significant differences in any of these tests. Thus, regardless of the food ratio, and regardless of whether the stimuli were paired with food or not, the overall left: right red keylight ratio had no effect on the overall behaviour ratio ($N = 5$, $-1.75 < z > -1.21$, $p > .05$ for all tests). In sum, no extended-level effects of the stimuli, whether paired or unpaired, or whether positively or negatively correlated with food, were detected in any of the analyses performed.

3.3.2 Preference trees

The log (left/right) response ratio after a food or a red keylight is presented in Figure 3.1 and 3.2 as a function of the number of preceding events of that sort (food or red keylight) from the same alternative (left or right). Figure 3.1 depicts these *continuation trees* for all 9:1 food ratio conditions in Phase 1 (unpaired stimuli; left panel) and Phase 2 (paired stimuli; right panel). Figure 3.2 depicts these continuation trees in all conditions arranging a 1:9 food ratio. Each log (left/right) response ratio was calculated for each subject in each condition (these individual plots are presented in Appendix Figures A1-A10). If there were fewer than 20 responses in total after a particular sequence, no response ratio was calculated for that individual. If fewer than 3 subjects had a valid data point at a particular sequence, the mean response ratio for that particular sequence was not calculated and is not shown. Figures 3.1 and 3.2 present the group mean log response ratios in each IRI.

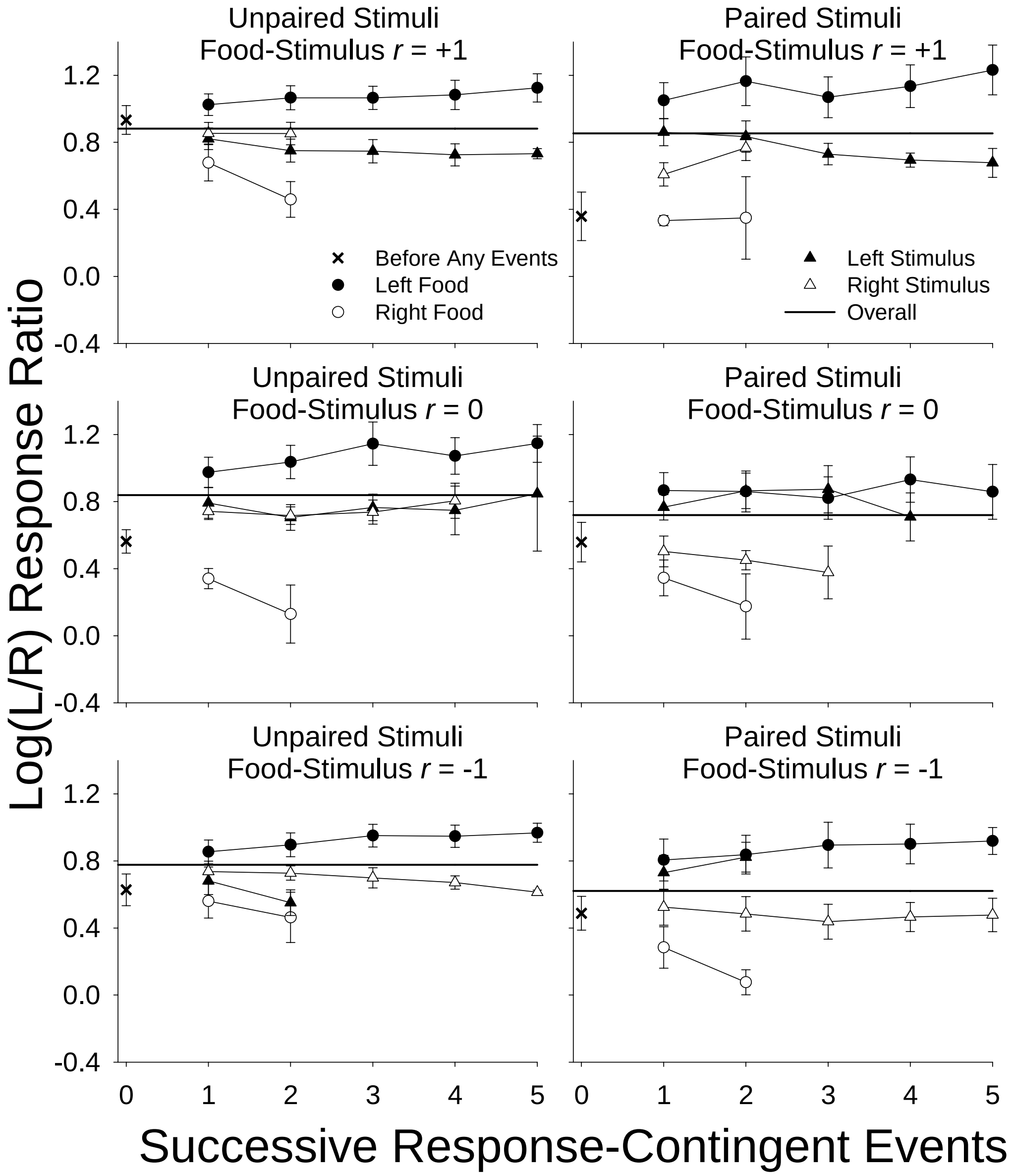


Figure 3.1. Group mean log (L/R) response ratio after a response-contingent event as a function of the number of preceding same-alternative same-events in the six conditions with response-contingent stimuli and a 9:1 food ratio. The stimuli were unpaired with food in the left panels and were paired with food in the right panels. Error bars depict one standard error.

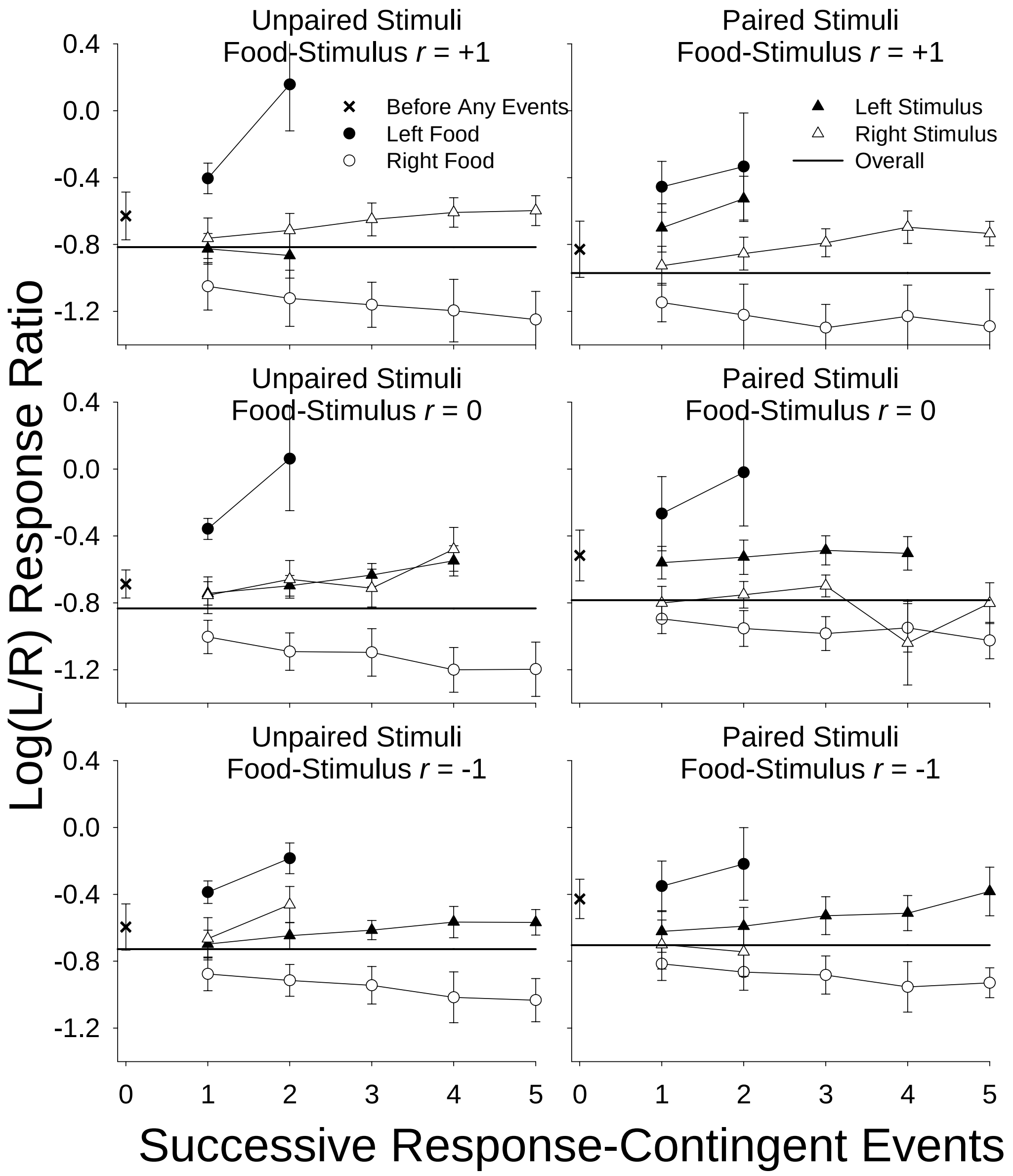


Figure 3.2. Group mean log (L/R) response ratio after a response-contingent event as a function of the number of preceding same-alternative same-events in the six conditions with response-contingent stimuli and a 1:9 food ratio. The stimuli were unpaired with food in the left panels and were paired with food in the right panels. Error bars depict one standard error.

Successive same-alternative foods have been found to bring preference progressively further towards the alternative providing those foods (e.g., Davison & Baum, 2000). Visual inspection of Figures 3.1 and 3.2 suggests that this was also true here. Nonparametric trend tests (Kendall) were conducted on the individual subject data (Appendix Figures A1-A10). In conducting these trend tests the value of *k* (the number of levels of the IV) was set at the lowest number of continuation reinforcers present for any animal. As the power of a trend test is reduced with fewer data points, only comparisons where $k > 2$ were considered. In 9 of the 12 cases that met this criterion (all of which were sequences from the overall richer alternative in that condition), the trend for preference to continually increase towards the alternative providing the foods was significant. There was a significant trend for preference to shift further towards the left as a function of successive left foods in Conditions 3, 18, 7, 4 and 16 and to shift further towards the right as a function of successive right foods in Conditions 5, 8, 6 and 17. Thus, there was a small, although not universal, trend for preference to continually increase as successive same-alternative foods were delivered. This conclusion is supported by visual inspection of the individual subject plots (Appendix Figures A1-A10) and is also consistent with previous findings obtained in steady-state environments (Landon et al., 2002).

A significant trend for preference to become more extreme as successive same-alternative keylights were delivered was found in 8 of the 16 trend tests where $k > 2$. Of these 8 significant trends, 3 occurred in conditions where the red keylights were paired with food (left stimuli in Conditions 18 and 17 and right stimuli in Condition 19), and 5 occurred in conditions where the red keylights were unpaired with food (left stimuli in Condition 6, right stimuli in Conditions 4 and 5 and both left and right stimuli in Condition 8). In 3 of the significant comparisons, the food ratio-stimulus ratio correlation was +1 (left stimuli in Condition 18 and right stimuli in Conditions 5 and 19), in 3 it was -1 (left stimuli in Conditions 6 and 17 and right stimuli in Condition 4), and in 2 it was 0 (left and right stimuli in Condition 8). Thus, any trend for preference to become more extreme with successive same-alternative red keylights was inconsistent. Additionally, there was no apparent tendency for preference to shift more depending either on whether the stimuli were paired with food, or on the direction of the food-stimulus correlation.

There is an alternative way to investigate the effects of response-contingent events using continuation data. In 123 of the 125 comparisons possible, preference after a left food was further to the left than was preference after a right food preceded by an equivalent number of same-alternative foods (Pigeon 66 in Conditions 17 and 19 provided the only 2 exceptions). Preference after a left stimulus can also be compared to preference after a right stimulus preceded by the same number of same-alternative stimuli. Wilcoxon matched-pairs signed-ranks tests were conducted to compare preference after a left stimulus with preference after a right stimulus preceded by an equivalent number of same-alternative stimuli. Separate tests were conducted in each condition. In 5 of 6 paired conditions, and 0 of 6 unpaired conditions, preference following one or more left red keylights was significantly different ($\alpha = .05$) from preference following an equivalent number of right red keylights (Condition 17 was the only paired condition where preference after a left stimulus was not significantly different from preference after a right stimulus). In all paired conditions in which a significant difference was found, preference after a left red keylight was significantly further towards the left than was preference after a right red keylight.

On the one hand, the trend tests indicated that the red keylights were never (conditional) reinforcers in the sense of continually increasing preference to the productive alternative. On the other hand, preference after a left red keylight differed from preference after a right red keylight only when the red keylights were paired with food. While the former finding is in agreement with Davison and Baum's (2006) conclusions, the latter finding is not. It may instead offer some support for a pairing hypothesis account of conditional reinforcement. First, significant differences were obtained in all but one paired condition and not in any unpaired conditions. Second, whenever the difference was significant, preference was always towards the just-productive alternative, even when the food ratio-stimulus ratio correlation in that condition (r = -1) would suggest that preference after a red keylight should be for the not-just productive alternative.

Wilcoxon matched-pairs *t*-tests were done to compare preference after a food with preference after a red keylight from the same alternative preceded by an equivalent number of the same event from the same alternative. In 22 of the 24 *t*-tests conducted (2 per condition, one comparing left-alternative events and the other comparing right-alternative events), preference was significantly further towards the just-productive alternative when food was the preceding event(s) than when red keylight was the preceding event(s). Preference after a left stimulus was (overall) not significantly different from preference after a left food in Conditions 16 and 19. Visual inspection of the individual-subject data suggests that the 2 nonsignificant differences were likely obtained because of the data one subject, or only 2 data points in the unexpected direction. Regardless of the reasons for the 2 nonsignificant findings, the 22 significant differences support the conclusion that follows from visual inspection of Figures 3.1 and 3.2 (and the individual subject plots in the Appendix): preference after a food was always more extreme than preference after a red keylight preceded by the same number of same-alternative same-events. Thus, although preference after a paired stimulus may have been toward the just-productive alternative, paired stimuli did not shift preference towards the just-productive alternative to the same degree as did foods.

Figures 3.3 and 3.4 (and the individual subject plots in Appendix Figures A11-A20) seek to determine whether a red keylight from the alternative that provided the immediately-prior food had effects comparable to a food continuation. Figure 3.3 shows food continuations in conditions arranging a 9:1 food ratio, as in Figure 3.1, but also includes response-contingent stimulus presentations from the alternative that provided the previous food. Figure 3.4 shows choice after food continuations and same-alternative red keylights in conditions arranging a 1:9 food ratio. Choice following these two event types is presented as a function of the number of same-alternative foods preceding either the food or the stimulus.

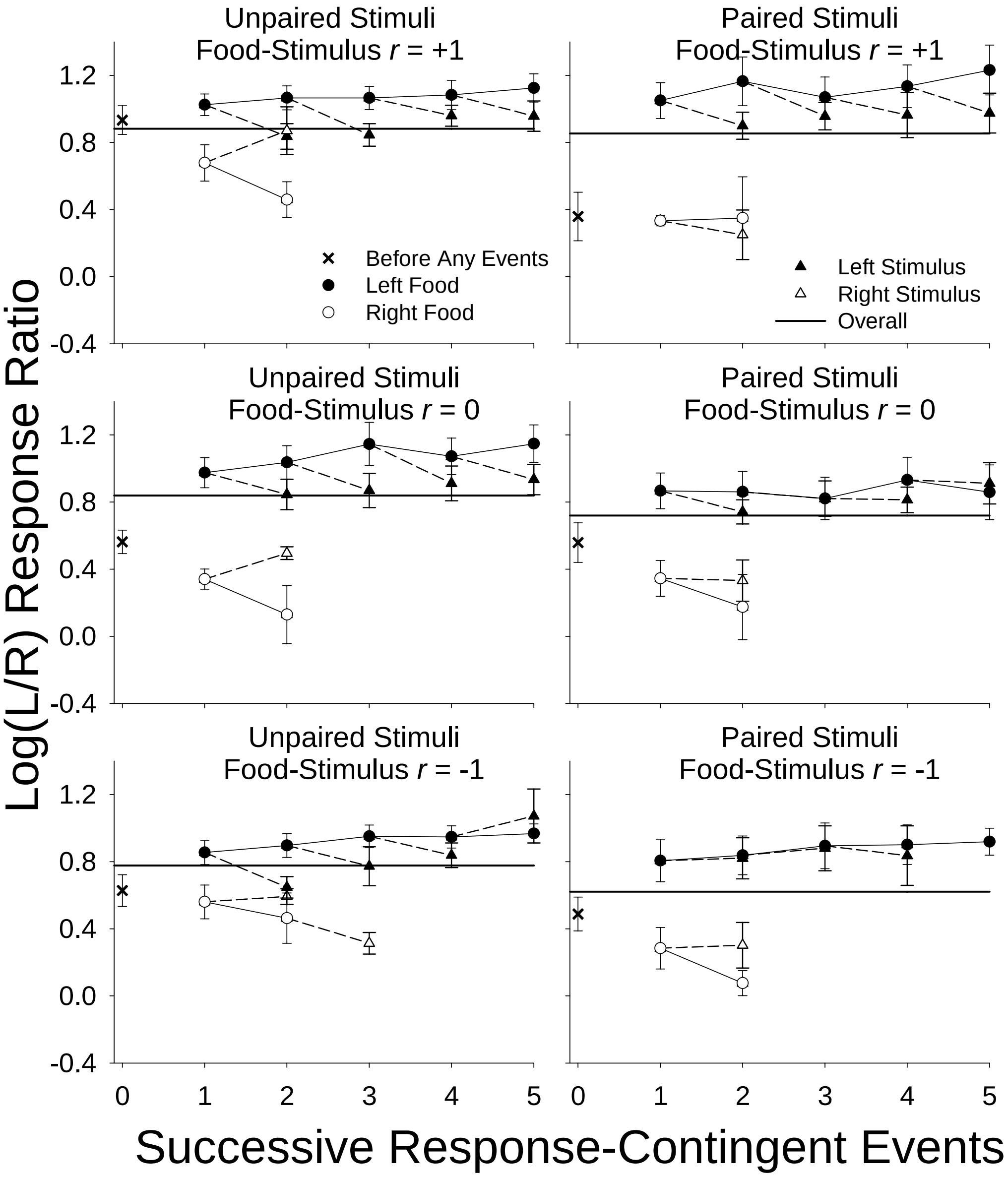


Figure 3.3. Group mean log (L/R) response ratio after one or more same-alternative foods (solid line), or a red keylight (dashed line) from the same alternative that delivered the prior food(s) in each of the six conditions with response-contingent stimuli and a 9:1 food ratio. The stimuli were unpaired with food in the left panels and were paired with food in the right panels. Error bars depict one standard error.

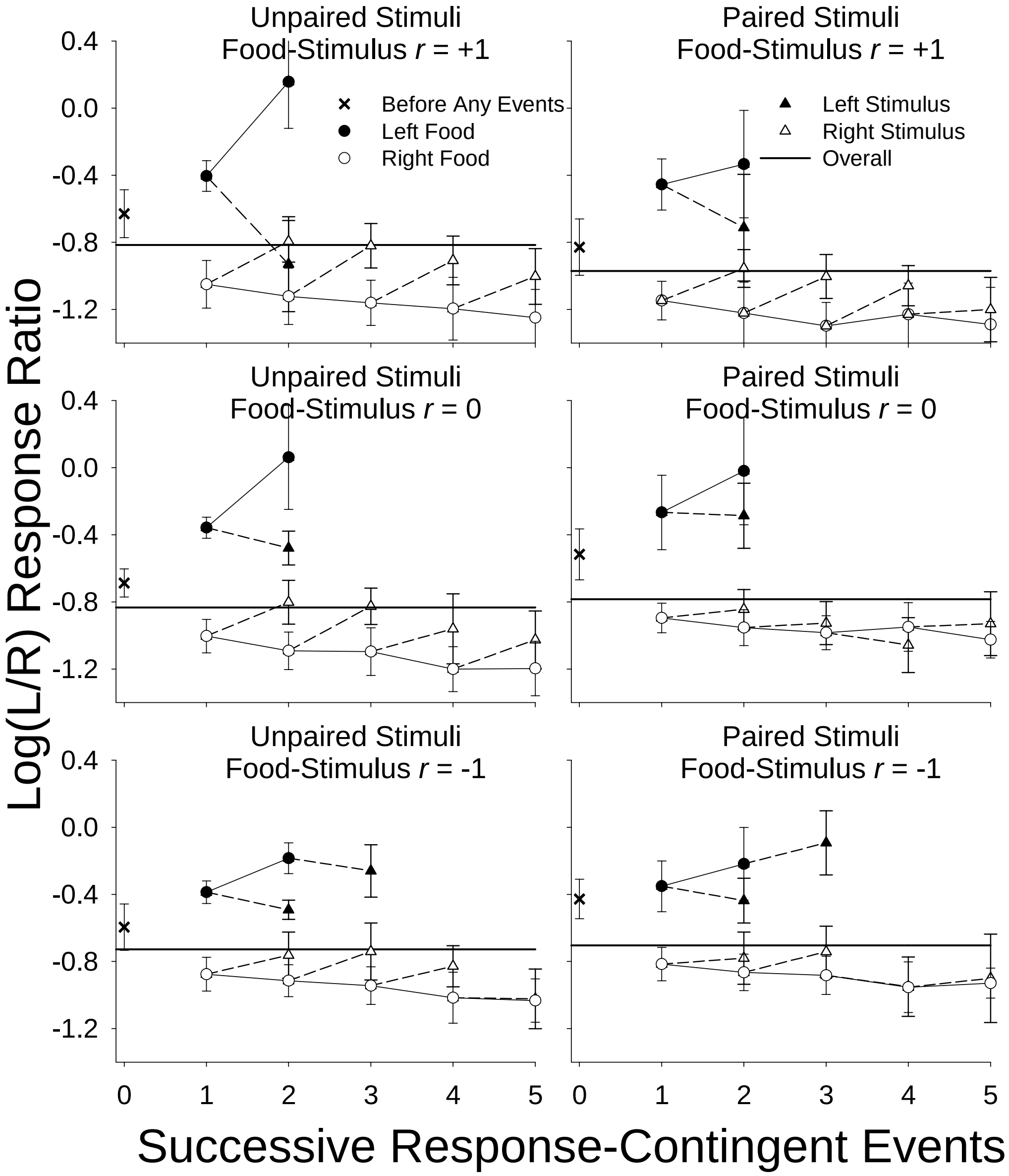


Figure 3.4. Group mean log (L/R) response ratio after one or more same-alternative foods (solid line), or a red keylight (dashed line) from the same alternative that delivered the prior food(s) in each of the six conditions with response-contingent stimuli and a 1:9 food ratio. The stimuli were unpaired with food in the left panels and were paired with food in the right panels. Error bars depict one standard error.

If the red keylights functioned as reinforcers, then preference after a red keylight from the same-alternative as the preceding food(s) would be equal to preference after a food continuation. Statistical tests indicated that this was not the case. Preference after a left food preceded by one or more left foods was significantly different (α = .05) from preference after a left stimulus delivered at the same point in 11 of 12 comparisons. Although the only nonsignificant result was in a paired condition (Conditions 16), for the most part, preference was always further towards the left after a left food than after a left stimulus. Preference after a right food preceded by one or more right foods was significantly further towards the right than was preference after a right stimulus preceded by the same number of right foods in 9 of 12 comparisons. Although 2 of the 3 tests in which preference after a right stimulus was not significantly different from preference after a right food were in paired conditions (Conditions 15 and 17), this is not compelling evidence that pairing made preference after a stimulus indistinguishable from preference after a food.

Discontinuations can also be investigated to determine whether the red keylights ever acted as apparent conditional reinforcers. A discontinuation is typically defined as a food from the alternative that did not provide the preceding series of foods. As there were both response-contingent foods and red keylight stimuli in this experiment, a discontinuation will be here defined as either a food or a stimulus from the alternative that did not deliver the previous food(s). Figures 3.5 and 3.6 (Appendix Figures A21-A30) show the same food continuations as Figures 3.1 and 3.2 but with added food or stimulus discontinuations.

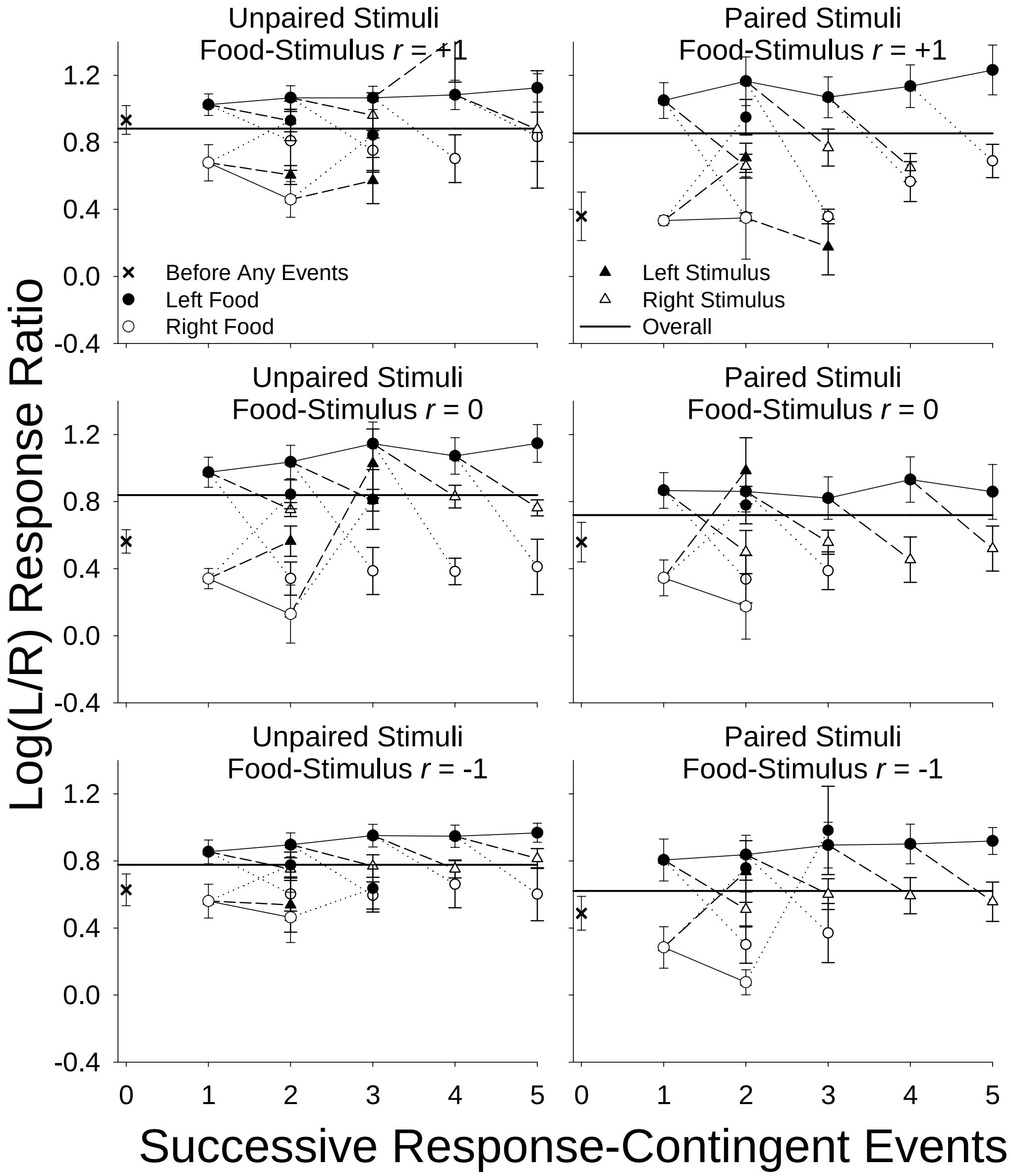


Figure 3.5. Group mean log (L/R) response ratio after one or more same-alternative foods (solid lines) or a food (dashed line) or red keylight (dotted line) that discontinued a sequence of same-alternative foods in each of the six conditions with response-contingent stimuli and a 9:1 food ratio. The stimuli were unpaired with food in the left panels and were paired with food in the right panels. Error bars depict one standard error.

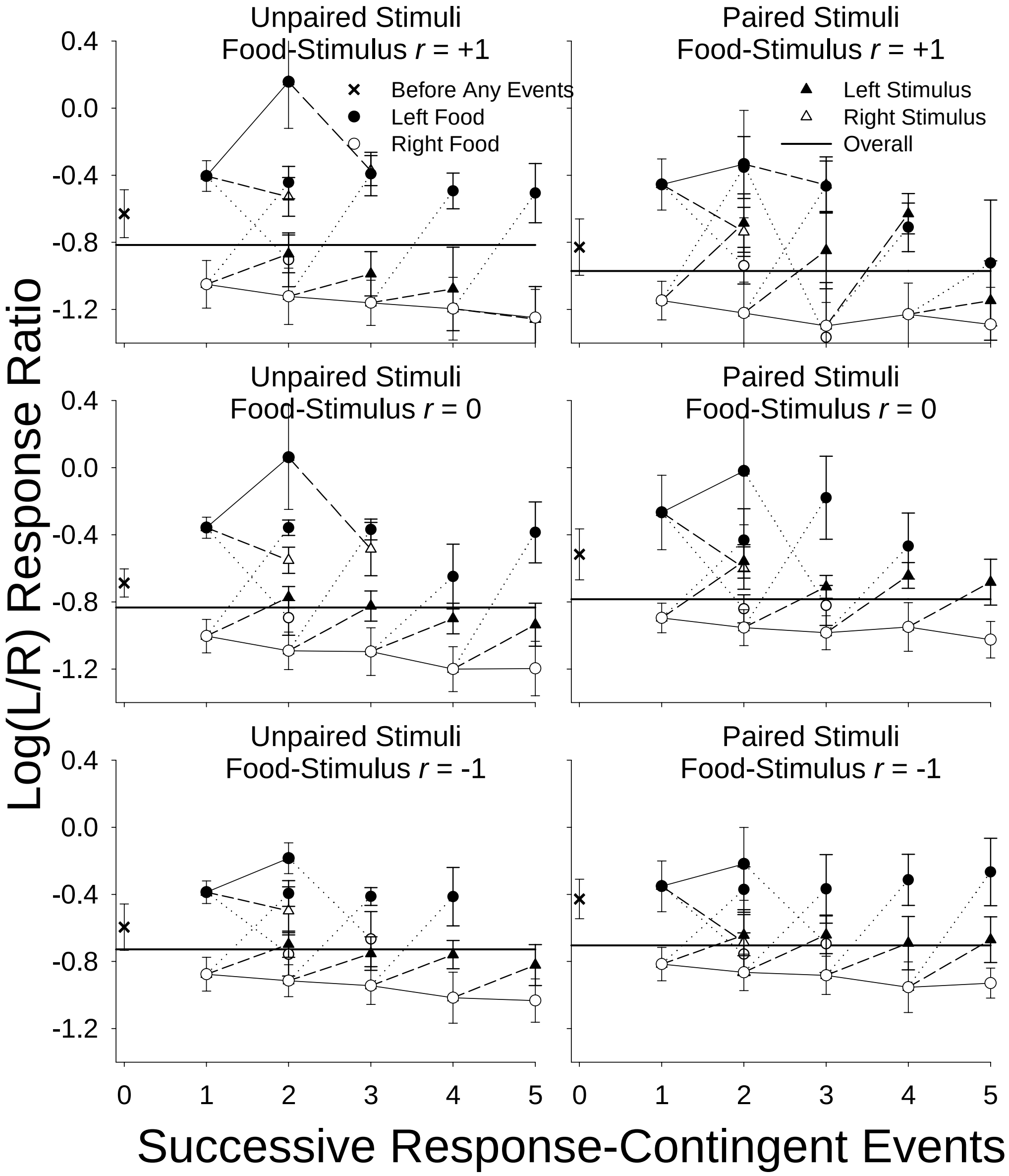


Figure 3.6. Group mean log (L/R) response ratio after one or more same-alternative foods (solid lines) or a food (dashed line) or red keylight (dotted line) that discontinued a sequence of same-alternative foods in each of the six conditions with response-contingent stimuli and a 1:9 food ratio. The stimuli were unpaired with food in the left panels and were paired with food in the right panels. Error bars depict one standard error.

Once again, there was no effect of the food ratio-stimulus ratio correlation in either the paired or the unpaired conditions in Figures 3.5 and 3.6 (or the individual subject plots; Appendix Figures A21-A30). There was also no apparent effect of pairing: within a single row, the plots did not differ from the unpaired to the paired condition. Wilcoxon *t*-tests conducted on the individual-subject data within each condition confirmed this. Preference after a food discontinuation was compared with preference after a stimulus discontinuation following an equivalent number of other-alternative foods. Preference was always significantly different ($\alpha = .05$) in all but one of the 24 comparisons. This indicates that preference was always shifted further by a food discontinuation than by a stimulus discontinuation; regardless of the food ratio-stimulus ratio correlation and regardless of whether the stimuli were paired or unpaired with food.

Neither forward-pairing the response-contingent red keylights with food, nor varying the food ratio-stimulus ratio correlation had any major effects on preference in the inter-event interval. While pairing may have created a preference for the alternative that delivered the last red keylight, there was no tendency for preference to continually increase as a function of successive same-alternative (paired or unpaired) red keylights (Figures 3.1 & 3.2). Preference after a (paired or unpaired) stimulus was also generally less extreme than was preference after a food reinforcer given a preceding sequence of either same-alternative (Figures 3.3 & 3.4), or other-alternative (Figures 3.5 & 3.6) foods. Thus, pairing had a very small effect at this level of analysis and the food ratio-stimulus ratio correlation had no effect whatsoever.

3.3.3 Preference pulses

The time within each session that each response and response-contingent event occurred was recorded. The log ratio of left to right responses in each 2-s time bin after a response-contingent event was calculated from these raw data. These log response ratios were plotted as a function of time since the event, and (for the group data) are shown in Figures 3.7 (9:1 food ratio conditions) and 3.8 (1:9 food ratio conditions). As in the tree figures (Figures 3.1-3.6), the left panels are conditions with unpaired red keylight stimuli and the right panels are conditions with paired red keylight stimuli. The top row shows conditions where the correlation between the left: right stimulus ratio and the left: right food ratio was +1, the middle row shows conditions where this correlation was 0 and the bottom row shows conditions where it was -1. Individual-subject preference pulses are presented in Appendix figures A31-A41. For each individual subject, if a time bin contained fewer than 20 responses in total, no data point was plotted. In cases where an individual subject's responding was exclusive to one alternative in a time bin, a log response ratio could not be calculated and a value of +4 or -4 was instead plotted. Only if a data point could be calculated for 3 or more individuals would a data point be plotted for the group. Error bars (which depict 1 standard error) are only plotted at representative data points at intervals of $\log_2$ time bins.

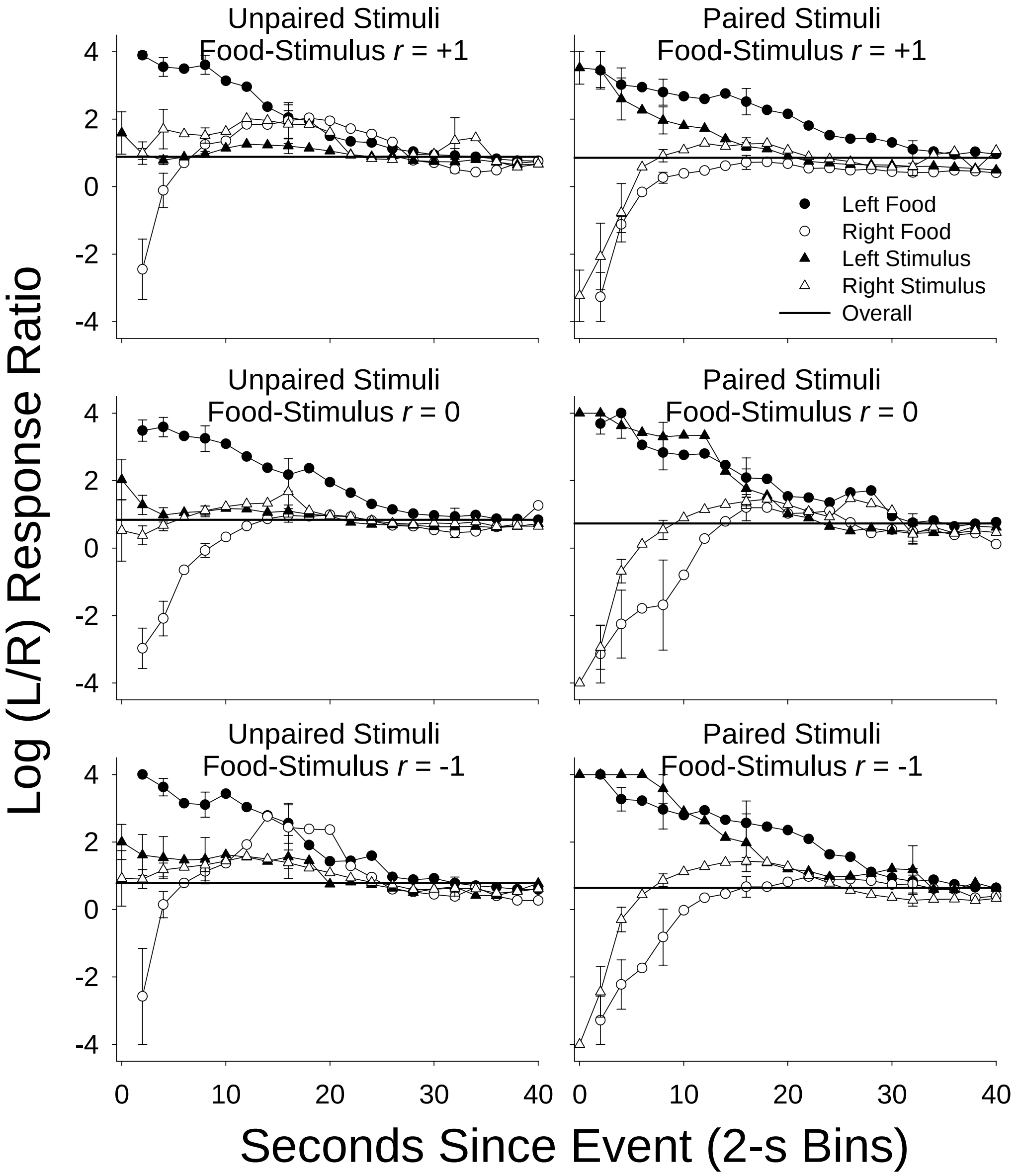


Figure 3.7. Group mean log (L/R) response ratio as a function of time (2-s bins) since the last response-contingent event in conditions arranging a 9:1 food ratio. Left panels depict conditions where the red keylights were unpaired with food and right panels depict conditions where the red keylights were paired with food. Error bars show ± one standard error and are plotted at representative data points at $\log_2$ time bins.

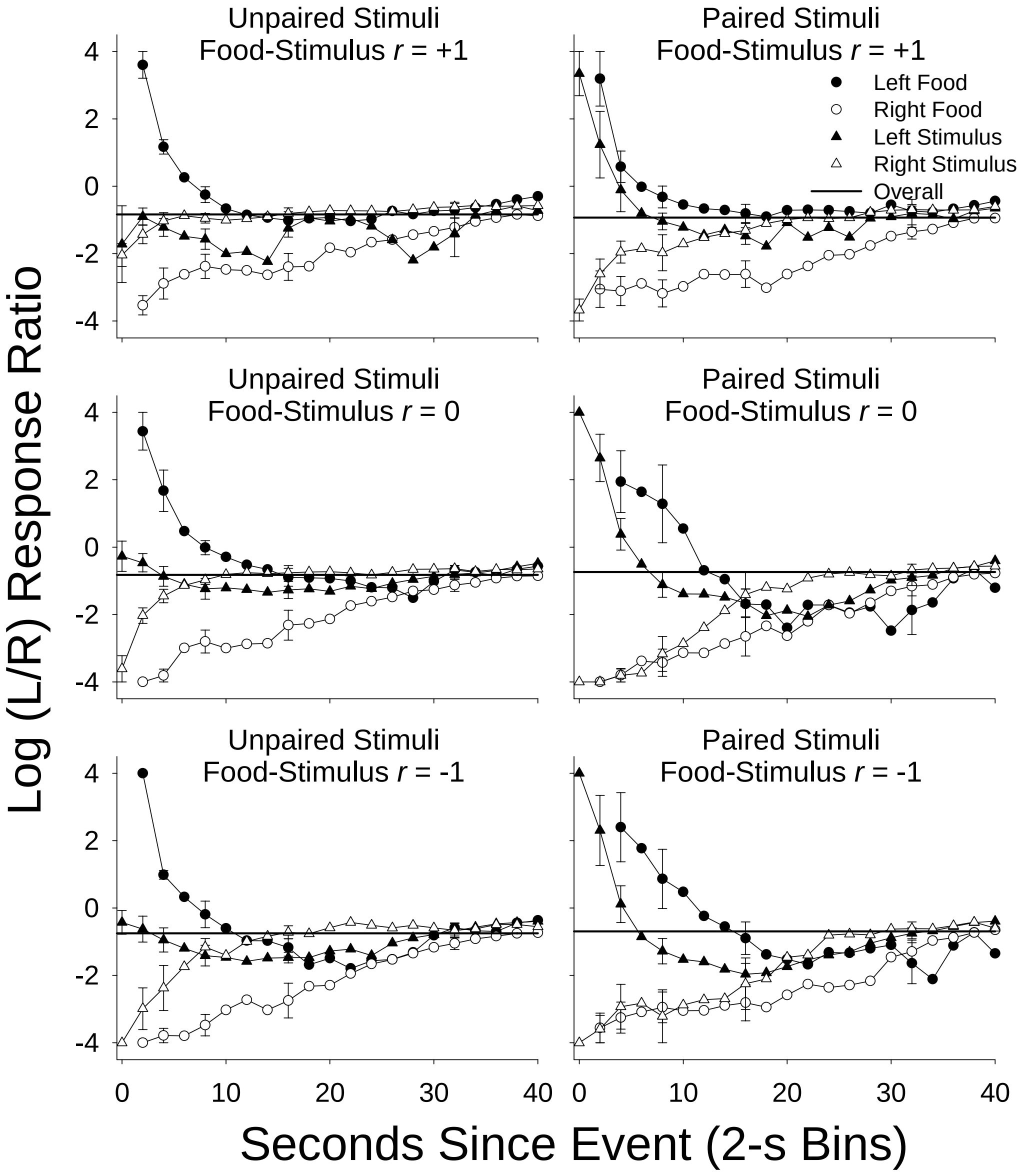


Figure 3.8. Group mean log (L/R) response ratio as a function of time (2-s bins) since the last response-contingent event in conditions arranging a 1:9 food ratio. Left panels depict conditions where the red keylights were unpaired with food and right panels depict conditions where the red keylights were paired with food. Error bars show ± one standard error and are plotted at representative data points at $\log_2$ time bins.

Generally, a response-contingent food was followed by a period of increased preference for the alternative which provided that food. This preference pulse lasted 20 to 35 s before falling to the extended preference level (solid lines). This finding is generally consistent with previously-reported post-food preference pulses. The post-stimulus preference pulses however were unlike those previously reported (Davison & Baum, 2006). Local preference after a red keylight stimulus was the same within a single column of Figure 3.7 and Figure 3.8. This indicates that the food ratio-stimulus ratio correlation had no effect on local preference immediately after a red keylight (as in Davison and Baum's study). There was however a discernable difference in the post-stimulus preference pulses going across a row. The post-stimulus preference pulses were all larger and further towards the just-productive alternative in the right columns (paired) than in the left columns (unpaired) of Figures 3.7 and 3.8. Pairing the red keylights with food increased preference to the just-productive alternative after that red keylight, and made the post-red keylight pulses more similar to the post-food pulses. This occurred regardless of the food ratio-stimulus ratio correlation. This pattern of results — no effect of the food ratio-stimulus ratio correlation, and a visible effect of pairing the stimuli with food — is inconsistent with Davison and Baum's (2006) results.

3.4 Discussion

The effects of varying whether a stimulus was paired or unpaired with food, and the correlation of the left: right stimulus ratio with the left: right food ratio were investigated in a steady-state concurrent schedule procedure. At the most local level of analysis (Figures 3.7 & 3.8), there was no detectable effect of the food ratio-stimulus ratio correlation. There was an apparent effect of stimulus-food pairing. Immediately after a paired stimulus, the log behaviour ratio reached extreme preference for the just-productive alternative, similar to the pattern seen after a response-contingent food. Preference pulses after unpaired red keylights were much smaller and could have been due to the changeover delay: a response-contingent event could not be delivered within 2 s of changing over from one alternative to the other. This effectively penalized switching at all times, including immediately after a response-contingent event — while food could be immediately arranged and obtained on the alternative that provided the just-delivered event, it could not be obtained on the other alternative for the duration of the changeover delay. CODs may generally contribute to all post-response-contingent event preference pulses (see Chapter 5) and may even be their sole cause, in particular when they are very small, as after the unpaired stimuli in the present experiment.

At a more extended level of analysis (preference trees), successive same-alternative foods continually shifted preference towards the alternative providing those foods, although to a lesser extent than in frequently-changing environments (Davison & Baum, 2000). The effects reported here were similar to those reported by Landon et al. (2002) who also arranged a constant food ratio. A series of stimulus continuations did not bring preference to levels nearly as extreme as did the same number of food continuations. Preference after a paired stimulus (but not an unpaired stimulus) was somewhat towards the just-productive alternative however. This finding, although reported in reference to the preference trees, merely quantifies the differences between paired and unpaired stimuli at the most local level of analysis (the preference pulses). It shows that preference after a paired left red keylight differed from preference after a paired right red keylight, but that there were no differences between unpaired left and right red keylights. Nothing is revealed about the effects of successive same-alternative red keylight deliveries or the effects of a red keylight after a series of same-alternative foods.

Preference after any response-contingent stimulus (paired or unpaired, positively, negative or zero-correlated with food) that followed one or more foods from the same alternative was always less extreme than was preference after a food continuation delivered at that point. Discontinuation foods brought preference further to the just-productive alternative than did discontinuation stimuli, and this was the case whether the stimuli were paired or unpaired, positively, negatively or zero-correlated with food. Thus, although the response-contingent stimuli, more specifically the paired response-contingent stimuli, did have some effect at the most local level of analysis, there was no real unique effect at this intermediate level of analysis. At the most extended level of analysis, there was no effect of pairing, correlation, or their interaction on global measures of sensitivity to the food ratio (Table 3.2). In sum, all response-contingent red keylights had a detectable effect at the most local level of analysis (paired keylights having the greater effect). At the more extended level of preference trees, red keylights always had greatly reduced effects (relative to food). There was no effect of any stimuli, paired or unpaired, positively, negatively, or zero correlated with food in the most extended-level analysis.

The present experiment's procedure was similar to that of Davison and Baum (2006): both paired and unpaired non-food response-contingent stimuli whose left: right delivery ratio could either be positively, negatively or zero correlated with the left: right food ratio, were inserted into a concurrent schedule of food reinforcement. The results of the two experiments were different however: Davison and Baum found that the food ratio-stimulus ratio correlation determined the direction of the post-stimulus preference pulse and that pairing had a very small effect. In the present study, the food ratio-stimulus ratio correlation had no demonstrable effect while pairing the stimulus with food did have an effect on post-stimulus preference.

Together, Experiment 1 and Davison and Baum's (2006) experiment might suggest that pairing a stimulus with food leads to a similarity in local food and stimulus effects only when environmental variation is low (as in the current experiment), and that positively correlating a stimulus with food leads to such a similarity only when environmental variation is high. While this may account for the failure to find an effect of correlation in the present experiment, it does not adequately explain the failure to find an effect of pairing in Davison and Baum's experiments. This difference might instead have resulted from the fact that different pairing procedures were used in each study. In the present experiment, the paired stimulus was a red keylight illumination which preceded food on paired trials. In Davison and Baum's study, the paired stimulus was a magazine-light illumination presented simultaneously with food on paired trials.

Within Pavlovian paradigms, simultaneously pairing a neutral stimulus with a primary reinforcer in order that the neutral stimulus become a CS is generally considered less effective than preparations where the CS precedes the US (Mackintosh, 1983). Rescorla (1980) however found greater acquisition with simultaneously rather than forward paired stimuli when both stimuli were hedonically neutral. This suggests that the more appetitive, and hence more salient, stimulus may attenuate conditioning in simultaneous pairing procedures because it is attended to almost exclusively, leaving little attention for the other stimulus. Davison and Baum's (2006) finding that choice after a magazine-light was in the direction indicated by the food versus magazine-light correlation suggests that attention was allocated to the magazine-light even when it was presented at the same time as food. If pairing was the only relevant factor in creating a conditional reinforcer, there seems to be no reason to claim that the simultaneous pairing used by Davison and Baum was necessarily inferior in creating stimuli with apparent conditional value. Some findings in fact suggest that Davison and Baum's paired stimuli would *more* easily acquire conditioned value than the paired stimuli used in Experiment 1. Classical conditioning may be enhanced when the CS and US are spatially contiguous (Ellins & von Kluge, 1990; Sullivan, 1984). Davison and Baum's CS (magazine light) was located on the apparatus which provided the food, while the stimuli used in the present experiment (keylight) were spatially distant.

Theories proposed to account for behaviour in concurrent chains (Grace, 1994; Killeen, 1982; Mazur, 2001; Squires & Fantino, 1971) may simply and parsimoniously account for the present findings. According to these theories, stimuli which signal a reduction in waiting time to food will (at least appear to) become conditional reinforcers — only the paired stimuli in Experiment 1 signalled such a reduction in waiting time to food. The unpaired stimuli in the present study, as well as both the paired and unpaired stimuli in Davison and Baum's experiment signalled no such reduction. Autoshaping research also predicts that the reduction in time to primary reinforcement signalled by a stimulus is related to that stimulus becoming a conditional reinforcer. When food reliably and noncontingently follows the illumination of a pecking key, a naïve pigeon will come to keypeck despite there being no requirement to do so (Brown & Jenkins, 1968) — and sometimes even despite a contingency that removes the food upon a key-peck (Williams & Williams, 1969; Woodard, Ballinger, & Bitterman, 1974). In this procedure, the CS will only reliably come to elicit key-pecking if the ratio of the CS-US interval and the US-US interval is sufficiently small (Gibbon & Balsam, 1981; Jenkins, Barnes, & Barrera, 1981). Learning in this case is also dependent on the CS signalling a reduction in the delay to reinforcement (relative to the average delay). Theories of Pavlovian conditioning (Rescorla & Wagner, 1972) also claim that a stimulus which provides novel and non-redundant information on a forthcoming US will become a CS. While the stimuli in Davison and Baum's experiment may have signalled which alternative was more likely to provide the next food, they did not signal that food was now temporally closer, thus accounting for their failure to find an effect of pairing. Only the paired stimuli in Experiment 1 signalled that the likelihood of food was now higher than it was prior to that stimulus delivery. This signalled increase in the local food rate may have caused the post-pairing greater preference to the just-productive alternative.

The present results also differed from those of Davison and Baum (2006) in that the effect of the food ratio-stimulus ratio correlation obtained in a frequently-changing environment was not found in a constant environment. In the frequently changing procedure, the extended reinforcer ratio does not predict the current-component reinforcer ratio. Each food is therefore a source of non-redundant information on the current food ratio, increasing the accuracy of the estimate of the current food ratio. Response-contingent stimulus deliveries, when they have a consistent relationship with food deliveries across the alternatives, can provide additional information about the unknown current contingencies.

In the present experiment on the other hand, the contingencies of food delivery were in place for periods of time long enough for behaviour to stabilize. Extended-level sensitivity to reinforcement was high in the absence of response-contingent stimuli. Any information about the left: right food ratio provided by the stimuli was redundant with information provided by the foods themselves. The addition of the stimuli, their pairing relationship with food, and their correlation with food made no difference to global measures of sensitivity (Table 3.2). There was simply no room for the stimuli to improve control.

The present absence of an effect of the food-stimulus correlation is consistent with a number of findings. In Pavlovian preparations, when a US is paired with two CS's, either the CS initially paired with US (blocking; Kamin, 1969) or the more salient CS (overshadowing; Pavlov, 1927) acquires greater ability to elicit the CR. It is also consistent with Bayesian foraging perspectives (McNamara, Green, & Olsson, 2006; Stephens & Krebs, 1986): In Davison and Baum's (2006) experiment, the probability of a left (vs. right) reinforcer was .5 prior to a stimulus. The posterior probability after a (positively or negatively correlated) stimulus differed from this prior probability. In the present experiment on the other hand, the prior probability of a left (vs. right) reinforcer was either .9 or .1 (depending on the reinforcer ratio in that condition) and this probability was unchanged by a response-contingent stimulus presentation.

Previously, the analytical tools and methods developed by Davison and Baum and colleagues in frequently-changing environments have been successfully applied to standard steady-state concurrent schedules (Landon et al., 2002, 2003b), suggesting that the mechanisms governing behaviour allocation are the same regardless of rate of environmental variation. However the present experiment demonstrated that this is not always the case. In a frequently-changing procedure, the reinforcer ratio is unknown at the start of a component, and the animal is likely to use all information on the likely location of future food. When however, the current food ratio is adequately signalled by the extended-level food ratio, the information provided by the additional response-contingent stimuli is redundant and these additional sources of information can be ignored. Whether added stimuli come to produce effects consistent with a conditional reinforcement interpretation depends on the context in which they are delivered. Two important aspects of context are: first, how the stimuli are paired with food; and second, the degree to which the contingencies are already discriminated.

Chapter IV

4.1 Experiment 2

Davison and Baum (2006) found that whether a response-contingent nonfood stimulus would come to act as if it were a conditional reinforcer, mimicking the local effects of a primary reinforcer, was determined by the correlation of the left: right stimulus ratio with the left: right food ratio. The stimulus' history of pairing with food had no effect. Precisely the opposite pattern of results was found in Experiment 1: the food ratio-stimulus ratio correlation had no visible effect on the preference pulses, while the stimulus' pairing history appeared to very strongly determine local preference after a red keylight.

The absence of an effect of the food ratio-stimulus ratio correlation in Experiment 1 (c.f., Davison and Baum, 2006) was attributed to differences in what the response-contingent stimuli signalled about the location of the next food in the steady-state (versus frequently-changing) environment. Due to the reinforcer ratio changing many times each session, sensitivity to the current reinforcer ratio is often rather low at the start of a component in frequently-changing environments, taking up to 8 reinforcer deliveries to reach an asymptotic level of about .4 -.7 (Davison & Baum, 2000; Landon & Davison, 2001). In frequently-changing environments, response-contingent stimuli can provide novel and non-redundant information about the current (unknown) food ratio (when the correlation between the food ratio and stimulus ratio is non-zero). In steady-state environments on the other hand, what the stimuli signal (the likely location of future foods) is redundant with what the animal already knows based on extended experience with the, unchanging, reinforcer ratio. In Experiment

1, group mean sensitivity to reinforcement was .94 in the absence of any response-contingent red keylights.

Theories of Pavlovian conditioning often cite stimulus information as crucial in determining whether a stimulus will become a CS (Rescorla, 1967; Wasserman & Miller, 1997). The non-redundancy or novelty of this information is also important to such theories (Kamin, 1969; Rescorla & Wagner, 1972). Thus, perhaps response-contingent non-food stimuli, such as those in Experiment 1 and Davison and Baum's (2006) study, produce apparent conditional reinforcer effects only when they provide novel, non-redundant information about the contingencies of food delivery. This was achieved in different ways in the two different experiments: in Experiment 1 pairing the stimuli with food led to an increase in their information-value: a paired stimulus signalled that the local food rate was now greater than in the keylight's absence (when the local food rate was zero). In Davison and Baum's experiment, a stimulus signalled the likely location of the next food, when the food ratio-stimulus ratio correlation was nonzero. In Experiment 2, *unpaired* stimuli non-redundantly signalled an increase in the local food rate, as the paired stimuli did in Experiment 1. This was achieved by arranging that response-contingent keylights and foods strictly alternated. Thus, a red keylight signalled that the next response-contingent event would be a food. The stimuli also (sometimes) signalled the local food ratio, as did the stimuli in Davison and Baum's experiment. In Experiment 2a, response-contingent changes in keylight colour (white to red) lasting 3 s (the same as the duration of food delivery) were interspersed between food deliveries in a steady-state concurrent schedule and the probability of a reinforcer on the same alternative as the last red keylight was varied. In Experiment 2b, the red keylight was extended to fill the entire interval until the next food. The stimuli were then reduced to their original duration (3 s) to evaluate the long-term effects of extending the stimuli. In Experiment 2c, additional stimuli (green keylights) were added (the red keylights remained). The local food ratio signalled by a green keylight differed from the local food ratio signalled by a red keylight.

4.2 Experiment 2a

Across the conditions of Experiment 2a, the probability that the next food would be on the same alternative as the last response-contingent red keylight (p_{same}), the overall food ratio, and the overall response-contingent red keylight ratio were all varied. When the overall food ratio was 1:1 and the probability that the next food would be on the same alternative as the last stimulus was either .9 or .1 (Conditions 2 and 3 respectively), stimuli from the left and right both provided unique and non-redundant information about the likely location of the next food. Prior to the stimulus delivery, the probability of a left food was .5. After the stimulus, this probability was either .9 or .1. In Condition 4 the overall left: right food ratio was 9:1 and the probability of a same-alternative reinforcer after a red keylight was .9. Thus, while the local left: right food ratio after a right stimulus (1:9) differed from the overall food ratio (9:1), the local food ratio after a left stimulus (9:1) was equal to the overall food ratio. In Condition 5 the overall left: right food ratio was 1:1, as was the local left: right food ratio after any keylight. Many more (90%) of these (uninformative) red keylights were on the left in Condition 5. In Condition 6, red keylights continued to be completely uninformative about the likely location of the next food and there were equal numbers of them on the left and right.

While the red keylights only signalled the next food’s likely location in Conditions 2-4, they always signalled an increase in the local food rate. This was also true of the paired stimuli in Experiment 1: food only appeared after a red keylight (though not after every red keylight). The red keylight stimuli in Experiment 2, unlike the stimuli in Experiment 1 which signalled an increase in the local food rate, were never contiguous with food. Food always followed a peck to a white keylight. Additionally, while the stimuli in Experiment 1 were never uniquely informative about the likely location of future foods, the stimuli in Experiment 2a did sometimes provide novel, non-redundant information about the likely location of future food (which Davison & Baum, 2006; 2010 found determined the direction of local post-stimulus preference). We may thus expect, as was demonstrated in Experiment 1, an effect of signalling that food is temporally closer. This effect, if obtained, cannot be attributed to a direct strengthening function of the response-contingent stimuli, acquired by virtue of the temporal contiguity of the red keylight and food. Second, an effect of signalling the likely location of the next food (as demonstrated by Davison and Baum, 2006; 2010) may be expected in Experiment 2a. This second effect should be absent in Conditions 5 and 6, when the local left: right food ratio after a red keylight was equal to the overall food ratio (1:1 in both cases). Condition 4 presents an interesting case: the overall left: right food ratio was equal to the local food ratio after a left stimulus (9:1 in both cases), but was different from the local left: right food ratio after a right stimulus (1:9).

Recent research suggests that there may be no effect of signalling a local food ratio which differs from the overall food ratio, either in Condition 4 or in any other condition. In a discrete trials concurrent schedule procedure, McLinn and Stephens (2006) reported that Blue Jays preferentially responded to the alternative which provided a greater proportion of the total reinforcers, rather than the alternative signalled more likely to provide the next reinforcer. When the richer alternative provided .75 of all reinforcers, 5 of the 6 subjects were always more likely to respond to that alternative, even when a stimulus indicated that the next food was more likely (p = .75) to be on the other alternative. Tracking the overall likelihood of events may be preferred over attending to additional, external stimulus cues. Thus, although the red keylights in this experiment may be a source of non-redundant, novel information on the likely location of the next food, behaviour may not reflect this and may instead adhere to the global food ratio.

4.2.1 Method

Subjects

Six homing pigeons, all experimentally naïve at the start of training and numbered 11 through 16 served as subjects. Pigeons were maintained at 85% ±15 g of their of their free-feeding body weights by postsession supplementary feedings of mixed grain when required. Water and grit were freely available in the home cages at all times. The home cages were situated in a room with about 80 other pigeons.

Apparatus

Each pigeon's home cage also served as its experimental chamber. Each cage measured 385 mm high, 370 mm wide and 385 mm deep. Three of the walls were constructed of metal sheets and the fourth wall and floor were metal bars. Two wooden perches were 60 mm above the floor. One of these was parallel to and the other at a right angle to the back wall. Three 20-mm diameter circular translucent response keys were on the right wall. These keys were 85 mm apart and 220 mm above the perches. The keys required a force exceeding approximately 0.1 N to register an effective response when illuminated. The food magazine was also on the right wall, 100 mm below the centre key. It measured 50-mm high by 50-mm wide and 40- mm deep. A food hopper, containing wheat, was situated behind the magazine and was raised and illuminated during food presentations. All experimental events were arranged and recorded on an IBM-PC compatible computer running MED-PC IV software which was in a room adjacent to the colony room.

Procedure

As the animals were experimentally naïve prior to this experiment, they were first magazine-trained. The duration that the food hopper was raised and the magazine illuminated was progressively decreased, and the interval between hopper presentations was progressively increased, until all birds were reliably eating as indicated by daily weighing.

Autoshaping was then used to get the pigeons to peck lighted keys. One of the three keylights was illuminated. If no response was made after 4 s, the hopper was raised and the magazine illuminated for 4 s. If an effective response was registered while the key was lit, the food was immediately delivered. After the food, there was a period of 5 s in which no foods were arranged and the keylights were all darkened. After this period, a variable time (VT) schedule started. When this schedule timed out, a key was again chosen to be illuminated. Initially, a VT 5 s was used and this was increased across sessions to a VT 20 s. Autoshaping sessions lasted 60 minutes or until 120 foods had been delivered, whichever happened first.

VI training was started when all six pigeons were reliably pecking. A VI 5-s schedule of reinforcement was initially arranged and the VI schedule was progressively increased across sessions until the terminal VI 54-s schedule was reached. The switching key procedure (more detail below) was introduced at the same time as the VI 5-s schedule. The centre key was a switching key that allowed the pigeons to switch from responding on one side key to responding on the other. A programming error was not detected until 41 sessions into Condition 1. The error consisted of the left key being illuminated after a right food. Responses in this period were not effective until the animal had made a changeover. This error proved to have rather long-lasting effects: after correcting the error in the program, preference following a right food was still noticeably affected. After 73 sessions in the corrected Condition 1 the food ratio was manipulated in attempt to address this problem: a 9:1 ratio was active for 33 sessions followed by a 1:9 ratio for 17 sessions. At this point, local preference appeared similar to results previously obtained (with different subjects). Condition 1 was then reintroduced. Only data from this final implementation of Condition 1 will be presented.

The concurrent schedule contained elements of a two-key concurrent schedule and elements of a switching-key concurrent schedule: Foods were arranged on the left and right keys but explicit changeover responses to the centre key were required to switch from one key to the other (a changeover ratio procedure, as used by Davison and Baum, 2006; 2010). At any one time, only the left or the right key was illuminated white along with the centre switching key (illuminated red). In order to switch, the animal had to make 4 responses to the centre key. The first response turned off the sidekey the animal had been pecking and the fourth response turned on the other key and turned off the switching key. The first peck to the newly illuminated sidekey turned the centre back on. Originally a 2-response switching requirement was arranged but this was increased to 4 responses as part of attempts to correct the abnormal behaviour caused by the programming error noted above.

Condition 1 was a concurrent schedule of food reinforcement with an overall VI 54-s schedule (i.e., concurrent VI 27 s VI 27 s). Every 1 second, the MED-PC program sampled a probability generator and decided whether or not to arrange a food. Once a food was arranged, it was allocated to the next effective left or right keypeck with a probability of .5.

In Conditions 2 to 6, response-contingent red keylights were inserted. This increased the overall event delivery schedule to VI 27 s but kept the overall food schedule at VI 54 s as in Condition 1. Foods and stimuli strictly alternated and these events had to be collected before the event timer would recommence. Across conditions the local food ratio signalled by the red keylights, the overall food ratio, and the overall stimulus ratio were varied (Table 4.1). Each session lasted 60 foods or 60 minutes, whichever came first. Each condition lasted at least 65 sessions.

Table 4.1

Sequence of conditions in Experiment 2a along with the overall food ratio, overall stimulus ratio and probability of a food on the same alternative as the last red keylight.

Condition	Overall Food Ratio	Overall Red Keylight Ratio	p_{same}
1	1:1	—	—
2	1:1	1:1	.9
3	1:1	1:1	.1
4	9:1	9:1	.9
5	1:1	9:1	.5
6	1:1	1:1	.5

4.2.2 Results

In previous work with similarly long-running conditions, 15 sessions has been identified as an adequate transition period (e.g. Landon et al., 2002; Experiment 1). Behaviour took longer to adjust in the present experiment, so data from the last 20 sessions of each 65-session condition were analysed. Justification for this is presented in Appendix B. The log (left/right) response ratio was calculated in each 2-s time bin after each response-contingent event for each individual subject. No individual subject log response ratio was calculated if there were fewer than 20 responses in total in a time bin. A value of ±3.5 was used if choice responding in a time bin was exclusive to one alternative. These individual subject preference pulses are presented in Appendix Figures C1-C6. Figure 4.1 presents the group mean preference pulses in Conditions 2 to 6. A data point was plotted only if there were valid data points for at least 2 of the 6 pigeons. Error bars depict 1 standard error and are plotted at representative points incrementing in $\log_2$ units.

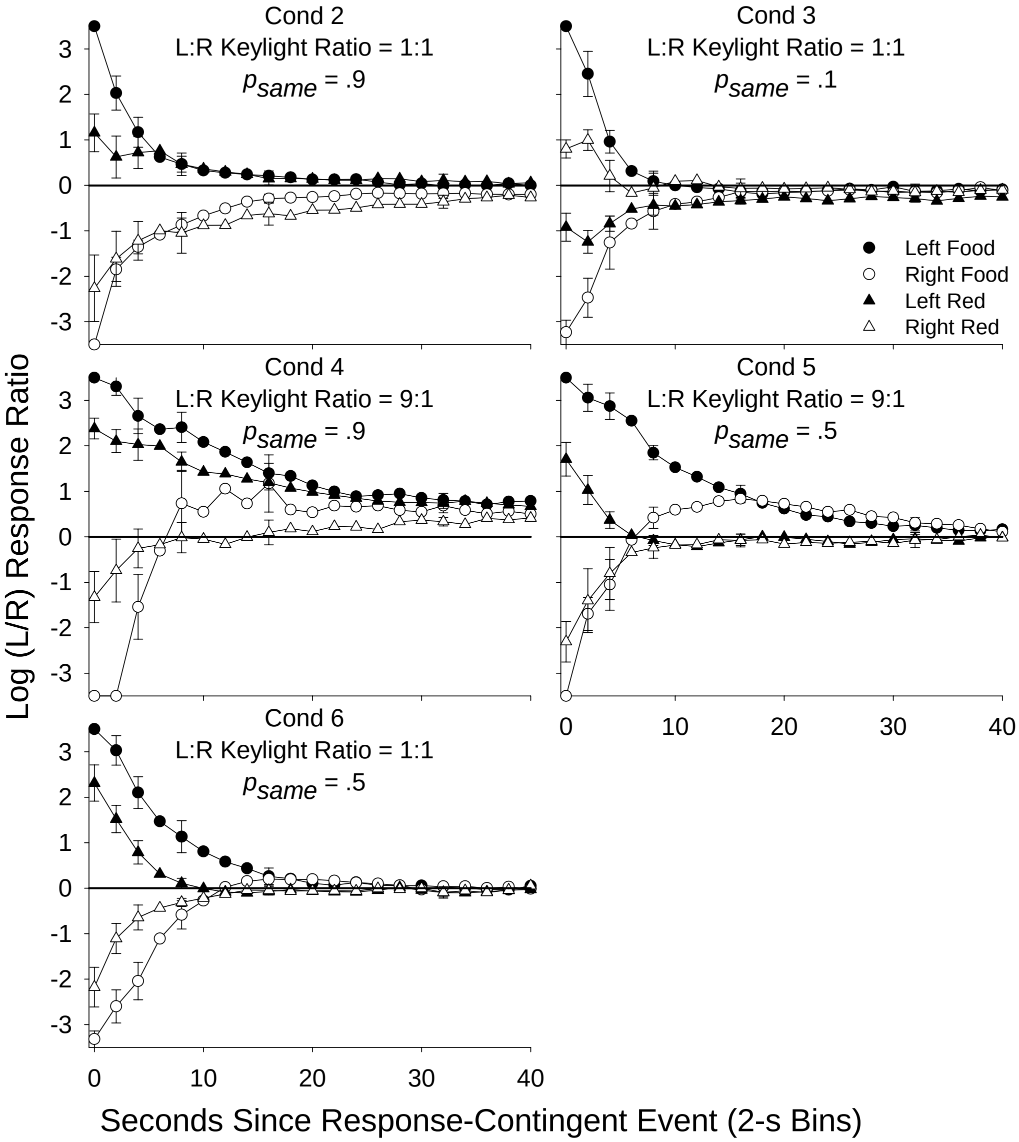


Figure 4.1. Group mean log (L/R) response ratio in each successive 2-s bin after each of the 4 response-contingent events in Experiment 2a. Error bars are plotted at representative data points and depict 1 standard error. The horizontal line is at 0.

If a local increase in preference to the just-productive alternative is interpreted as a typical, definitional reinforcer effect (an increase in the response previously followed by reinforcement; Skinner, 1938), then Figure 4.1 indicates that the red keylights were (conditional) reinforcers in Conditions 2, 4, 5 and 6: preference immediately after a response-contingent red keylight in those conditions was towards the just-productive alternative. In Conditions 2 and 4 the red keylights signalled that food was forthcoming and that this next food was very likely (p = .9) to be on the just-productive alternative. In Conditions 5 and 6, the stimuli *only* signalled that food was temporally closer. The preference pulses that followed these uninformative (about the location of the next food) stimuli suggest that simply signalling an increase in the local food rate from its prior value of 0 is enough to produce a transient increase in preference to the just-productive alternative.

There was also evidence of an effect of signalling the next food's likely location. First, preference after a red keylight in Conditions 5 and 6 fell to indifference within 15 s of the red keylight delivery. In contrast, preference after a red keylight in Condition 2 remained different from overall preference even 40 s after that red keylight. Second, preference after a red keylight in Condition 3 was towards the not-just-productive alternative, consistent with the low probability of a same-alternative reinforcer in that condition.

The preference pulses in Conditions 4 and 5 also indicate some control by the local food ratios. Although the overall log response ratio in Condition 4 was 0.83 (reflecting the 9:1 overall food ratio), choice for at least 40 s after a right red keylight was clearly at a level of (relative) preference for the right, reflecting the *local* 1:9 food ratio in that period. In Condition 5, preference after (left and right) red keylights started at the just-productive alternative before settling at approximate indifference. Preference after foods in that condition also started at the just-productive alternative but did not settle at indifference. Preference instead approached relative preference for the left. These differences in the level at which preference stabilized after the initial pulse to the just-productive alternative likely reflect the different arranged response-contingent event ratios in those periods: after a red keylight (from either the left or right), the left: right food ratio was 1:1, and after a food (from either alternative) the left: right red keylight ratio was 9:1.

Given that preference after any response-contingent event in Condition 5 stabilized at a level indicative of the local food ratio in that period, why did post-stimulus preference pulses in Conditions 2 and 3 approach approximate indifference (the overall food ratio) rather than a level reflective of the local food ratio? One possibility is that, as time since a food delivery increased, the *obtained* local food ratio approached indifference. This was investigated by calculating the log obtained response-contingent event (food and red keylight) ratios over time following a response-contingent event. The group means are presented in Figures 4.2 for Conditions 2 through 4, with the individual subject plots in Appendix Figures C7-C12. Left panels present the log (left/right) response and obtained red keylight ratios in each 2-s time bin after a food, and right panels present the log (left/right) response and obtained food ratios in each 2-s time bin after a red keylight. The individual-subject log obtained contingent event ratios across time were calculated in the same way as were the log response ratios across time with some small changes. The minimum number of response-contingent events per time bin was 5 rather than 20, and when the event ratio in a particular time bin was exclusive, a value of ±2 was plotted (rather than ±3.5 in the response analysis).

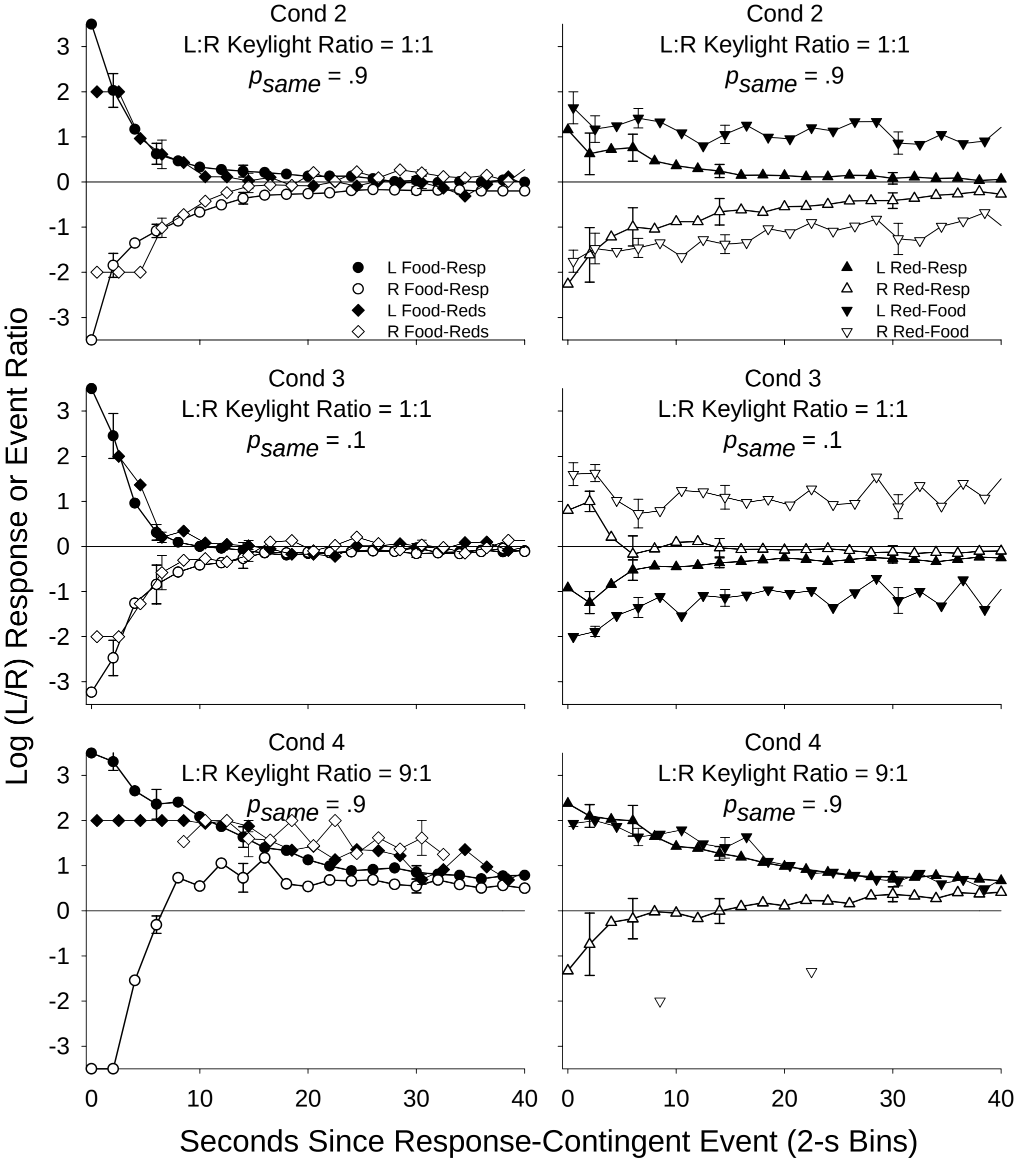


Figure 4.2. Group mean log (L/R) response and response-contingent event ratio throughout the post-event period in Conditions 2, 3 and 4. The left panel shows the response and contingent event pulses after a left or right food. The right panel shows these pulses after a left or right red keylight. Error bars depict one standard error and are plotted at representative data points that increment in units of $\log_2$. The contingent event pulses are shifted to the right of the response pulses by .5 s on the x-axis and are connected with a thinner line in order to aid comparison of the plots.

Preference after a food in Conditions 2, 3 and 4 (left panels) was strongly towards the just-productive alternative, as was the local obtained red keylight ratio in the same period. These extreme local obtained red keylight ratios, which appear inconsistent with the arranged local and global red keylight ratio (1:1 in Conditions 2 & 3), were likely caused by the extreme preference to the just-productive alternative immediately after a food. In the first 2 s after a food, preference was exclusive to the just-productive alternative. Preference approached indifference across time, but continued to be somewhat towards the just-productive alternative for at least 30 s. This extreme local behaviour ratio immediately after a food was likely initially due to the changeover contingencies. Once established however, it entered into a dynamical system (Davison, 1998) with the local obtained response-contingent event ratio. The extreme local behaviour ratio drove the local obtained food ratio which then further drove (reinforced) the extreme behaviour ratio (Herrnstein, 1970).

No such dynamical interaction was apparent between the local response and food ratios after red keylights in Conditions 2 and 3 (top and middle right panels). The local response ratio throughout the post-red keylight period was noticeably different from the local obtained response-contingent event ratio in this period. The local behaviour ratio after a red keylight in Condition 2 started at (somewhat) extreme preference for the just-productive alternative but shifted towards indifference through time. The local food ratio throughout this period however was consistently towards the just-productive alternative. The local obtained food ratios also started and remained further from indifference than the local behaviour ratios in Condition 3, although the deviations were towards the not-just-productive alternative. Any dynamical interactions between local behaviour and response-contingent event ratios after red keylights were not as prevalent as they were after foods.

Because of the more extreme arranged response-contingent event ratios in Condition 4, fewer local obtained food and red keylight ratios could be calculated, leading to more missing data points. After a left red keylight, both the log response and food ratios descended towards a level reflecting the 9:1 local food ratio (which in this case was equal to the overall food ratio). Although the local obtained food ratio after a right red keylight could only occasionally be calculated (due to a paucity of data), it too was clearly similar to the arranged local food ratio in that period (1:9). The local *behaviour* ratio after a right red keylight however clearly and consistently differed from this local (obtained and arranged) food ratio: Preference after a right red keylight in Condition 4 appeared to stabilize at the overall, rather than the local, food ratio. This suggests that the similarity of log response and obtained food ratios after left foods in Condition 4 was simply due to the local food ratio in this period being equivalent to the overall food ratio. As in Conditions 2 and 3, preference after the initial pulse to the alternative momentarily more likely to provide the next food descended towards the global rather than the local food ratio. Figure 4.3 shows the group mean preference and obtained response-contingent event pulses in Conditions 5 and 6 (the individual subject plots are presented in Appendix Figures C13-C18).

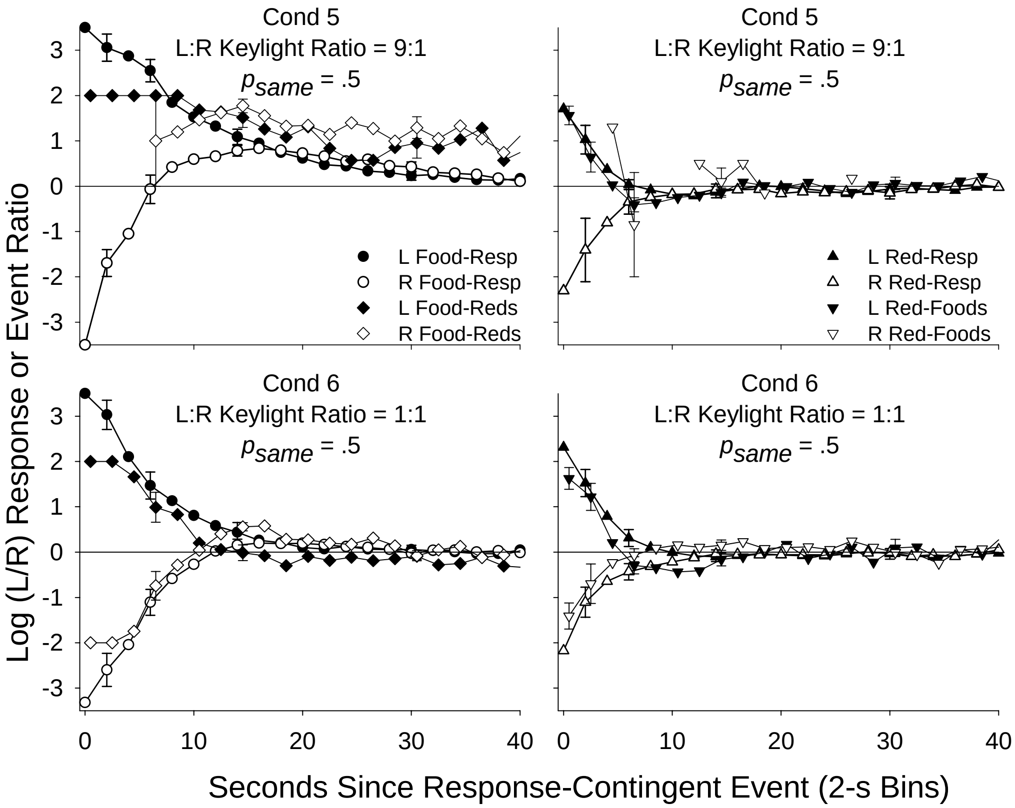


Figure 4.3. Group mean log (L/R) response and response-contingent event ratios throughout the post-event period in Conditions 5 and 6. The left panel shows the response and contingent event pulses after a left or right food. The right panel shows these pulses after a left or right red keylight. Error bars depict one standard error and are plotted at representative data points that increment in units of $\log_2$. The contingent event pulses are shifted to the right of the response pulses by .5 s on the x-axis and are connected with a thinner line in order to aid comparison of the plots.

In Conditions 5 and 6, as in Conditions 2-4, preference after a food closely followed the local obtained red keylight ratio. Unlike the earlier conditions however, preference after red keylights also closely followed the local obtained food ratio. This was likely because, in Conditions 5 and 6, the local food ratio after a red keylight (1:1) was equal to the overall food ratio, as it was after a left red keylight in Condition 4.

After the initial pulse to the just-productive alternative, choice after any food in Condition 5 stabilized not at indifference, but at a level of (some) preference for the left. Preference after red keylights however stabilized at indifference. The level at which preference stabilized was apparently determined by the local ratio of obtained response-contingent events in the period after that particular event type, rather than by the extended response-contingent event ratio. After a food (from either the left or right), preference tended toward a 9:1 ratio, and after a red keylight to a 1:1 ratio. Thus, although there was apparently greater control by the global than the local food and red keylight ratios, there was no apparent blending of these ratios.

The above data can be used to calculate local sensitivity to reinforcer values. These (group mean) sensitivity values are presented in Figure 4.4 for Conditions 2 and 3. Similar plots are not presented for Condition 6 because the arranged food ratio after left and right stimuli was 1:1 and the obtained values were close to this. Thus, estimates of sensitivity to reinforcement were highly variable and small changes in the local behaviour ratio had large effects on the calculated sensitivity values. Similar plots could not be created for Conditions 4 and 5 as there was insufficient data in a number of time bins.

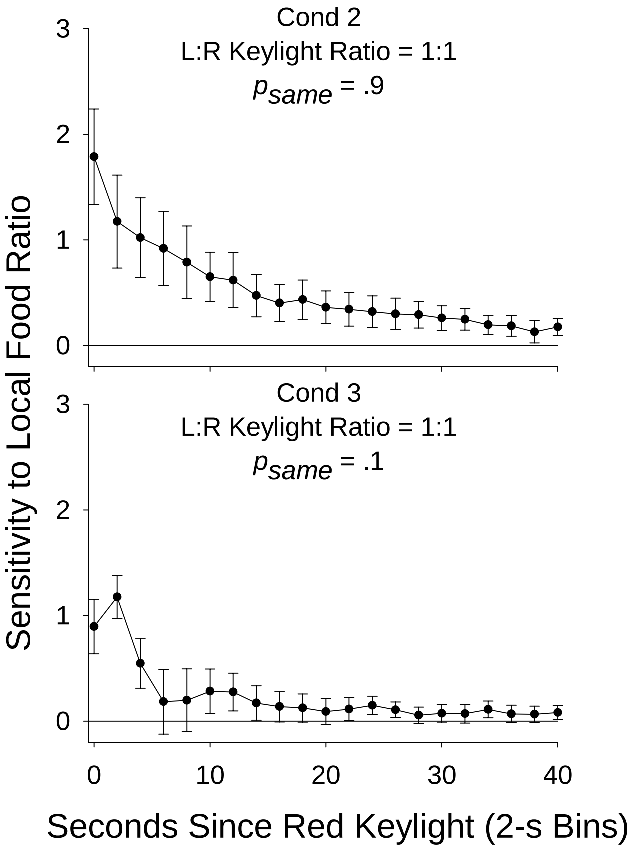


Figure 4.4. Group mean sensitivity of the local behaviour ratio in each 2 s bin after a response-contingent red keylight to the obtained food ratio in that period in Conditions 2 and 3. Error bars depict 1 standard error.

Figure 4.4 confirms conclusions from Figure 4.2: Sensitivity to the local obtained food ratio started higher and remained higher in Condition 2 than in Condition 3. Second, and also consistent with Figure 4.2, sensitivity to the local food ratio after a red keylight in Conditions 2 and 3 decreased throughout the post red keylight period.

4.2.3 Discussion

The local behaviour ratio after a red keylight only somewhat approximated the local food ratio when the local food ratio differed from the overall food ratio (after all red keylights in Conditions 2 and 3 and after a right red keylight in Condition 4).

When however the local food ratio was equal to the overall food ratio (after all red keylights in Conditions 5 and 6 and after a left red keylight in Condition 4), there was greater agreement between the local behaviour and obtained food ratios. These results suggest relatively weak control by the local contingencies of reinforcement as signalled by a red keylight, and relatively stronger control by the global food ratio. This result is similar to that reported by McLinn and Stephens (2006).

The apparently weak control by the local contingencies of reinforcement is surprising given a number of findings showing that behaviour can be highly responsive to often-changing local contingencies of reinforcement. Behaviour has been shown to vary with contingencies of reinforcement that change every six days (Davison & Hunter, 1979), every single day (Hunter & Davison, 1985), every 4 to 12 reinforcer deliveries (Davison & Baum, 2000), and arguably even after a single reinforcer delivery (Krägeloh et al., 2005). These previous reports of strong control by often-changing contingencies make the weak control by the local reinforcer ratios in the present experiment all the more perplexing.

In the frequently-changing procedure, sensitivity to the current component reinforcer ratio generally starts low and increases as a function of successive reinforcers. This increase in sensitivity to the current component reinforcer ratio is accompanied by a decrease in sensitivity to the previous component reinforcer ratio (Davison & Baum, 2000). In the present experiment, a food's influence on choice may have extended into the following post-red keylight period, attenuating control by that red keylight. To investigate whether control by the previous reinforcer interfered with control by the current stimulus, preference pulses after red keylights were plotted separately for red keylights preceded by a same-alternative food and for red keylights preceded by a food from the other alternative. These individual subject post-stimulus preference pulses (in Conditions 2 through 6) are presented in Appendix Figures C19-C24, and the group mean plots are presented in Figure 4.5. These preference pulses were created in the same way as were the earlier preference pulses with the exception that behaviour was further categorized according to the location of the prior food.

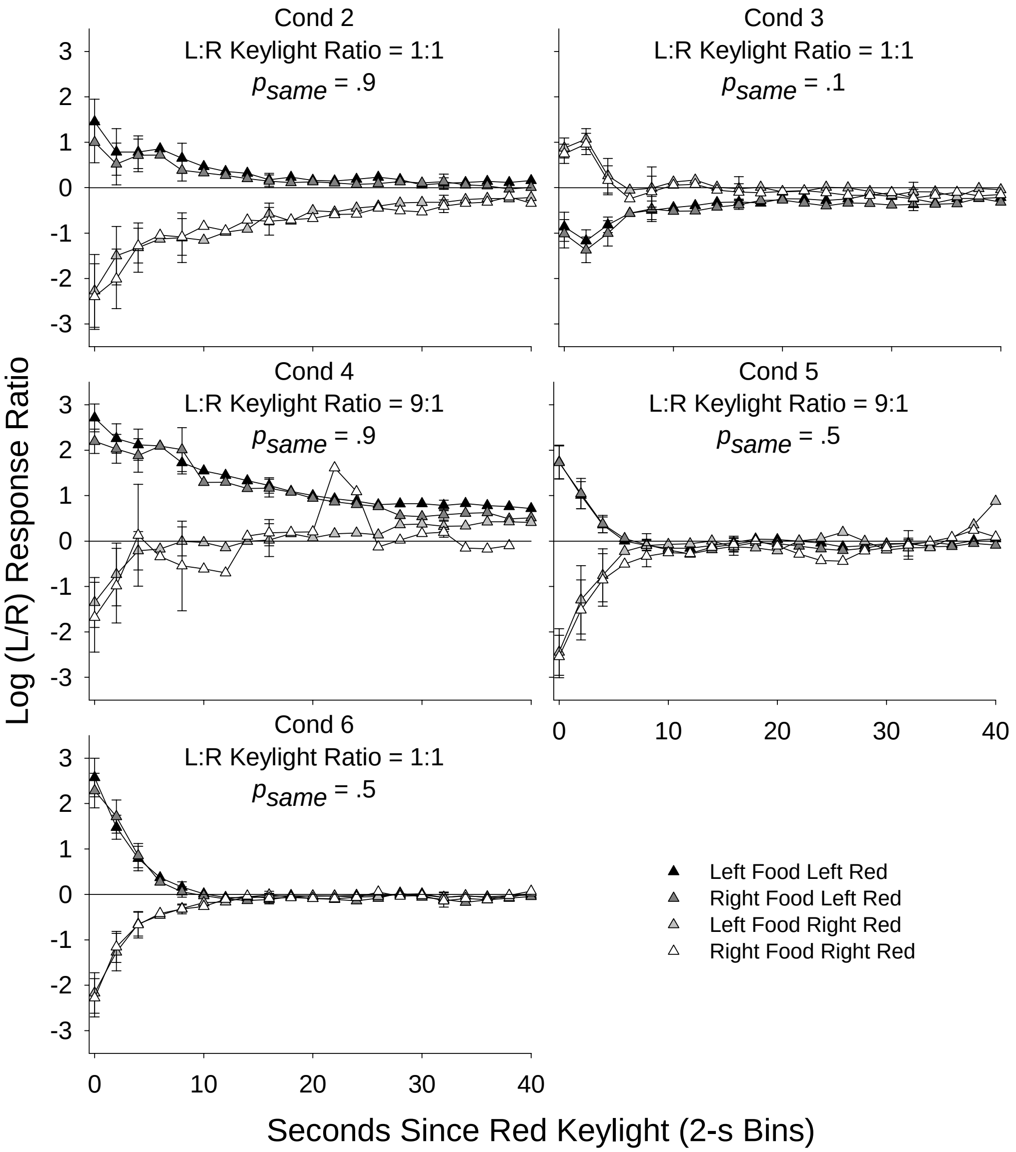


Figure 4.5. Group mean log (L/R) response ratio in each successive 2-s bin after left and right red keylights preceded by left and right foods in Conditions 2-6. Error bars are plotted at representative data points incrementing in units of $\log_2$ and depict 1 standard error.

Preference was clearly a function of the location of the last red keylight and there was no apparent influence of the prior food: the preference pulse after a left red keylight was the same whether the prior food was from the left or the right. The same was true for local preference after a right red keylight. Red keylights thus eliminated any control by or influence of the previous reinforcer. This finding is consistent with that of Krägeloh and Davison (2003) who found near-zero estimates of sensitivity to the previous component reinforcer ratio when distinct discriminative stimuli signalled the current component reinforcer ratio. Excess residual influence of the prior reinforcer was thus not responsible for the low agreement between local response and obtained food ratios.

Alternatively, failure of working memory may account for the low sensitivity to the obtained local food ratio at relatively long times from the last red keylight. The decrease in sensitivity to the local food ratio throughout the post-red keylight period (Figure 4.4) resembles a typical memory decay function (White, 2001; Wright, 2007). In memory research, the decrement in differential responding throughout the retention interval has been attributed to prior stimuli interfering with control by the most recent stimulus (Grant, 1975; Grant & Roberts, 1973). In these *trace strength theories*, memory of the sample stimulus is said to decay through time in a negatively accelerated fashion: immediately after a stimulus, the memory trace is strong relative to the trace of previously presented stimuli. Through time, the trace decays, becoming more similar to the trace from earlier stimuli (which are decaying at a slower rate). This leads to increasingly more interference, and thus a greater probability of an incorrect response throughout the retention interval.

Olton (1978) and Roberts and Grant (1976) both reported that accuracy earlier in a session was higher than accuracy later in the session. This finding supports the argument that proactive interference from earlier trials inhibits sample recall in later trials. Preference pulses after the red keylights were thus plotted separately for each quarter of the session (individual subject plots presented in Appendix Figures C25-C30). The log (left/right) response ratio in each 2-s time bin following a response-contingent red keylight was calculated for the first 15 red keylights in the session, as well as for keylights 16-30, 31-45 and 46-60. The group mean plots are presented in Figure 4.6.

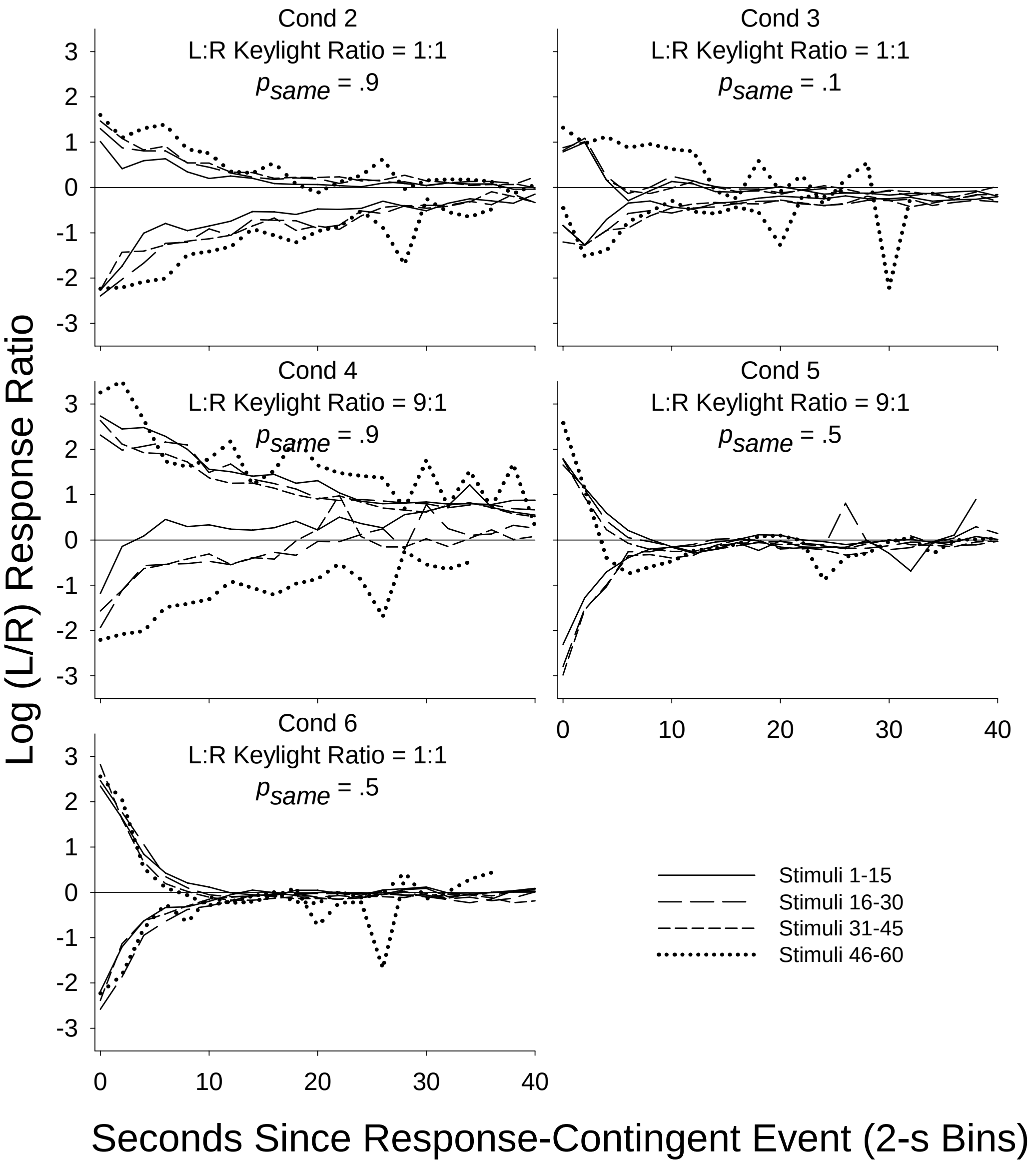


Figure 4.6. Group mean log (L/R) response ratio in each successive 2-s time bin after left and right red keylights in each of the 4 quarters of a session in Conditions 2 to 6. Note that sessions sometimes terminated before all 60 stimuli could be delivered, leading to greater variability in the plot for Stimuli 46-60 (dotted line) as well as the fact that only one such line is plotted in Condition 5.

There is no indication in Figure 4.6 that preference became less extreme throughout the session. If anything, there appears to be a tendency for preference after a red keylight to become *more* extreme later in the session in conditions with informative stimuli (Conditions 2 to 4). This may indicate some learning of what the red keylights signalled about the local contingencies of reinforcement throughout the course of each session, implying some failure of procedural memory between sessions. There was no evidence however of a failure of working memory: proactive interference from earlier red keylights did not contribute to the relatively low local sensitivity to reinforcement reported above (Figure 4.4).

4.3 Experiment 2b

Proactive interference is far from the only explanation proposed to account for memory decay. Honig (1978) for example, proposed a "memory as instruction" account in which the sample stimulus establishes an instruction about the behaviour that is to (eventually) follow. In the case of Experiment 2a, such an instruction might be in the form of "peck left" or "peck right". According to Honig, this instruction does not decay through time, rather it is either remembered completely (perhaps via some process of rehearsal) or forgotten completely. As support for this, Honig noted high within-subject variability in the duration over which correct responding could be maintained. A memory trace which decays at a constant rate cannot produce such high trial-to-trial variability. Although the progressive decrease in sensitivity to the local food ratio may not have been due to proactive interference, a failure of working memory may still have been responsible.

In Conditions 7 and 8, the red keylight stimuli were extended until the next food, thus eliminating any requirement to hold the location of the last red keylight in working memory. In all other respects, Conditions 7 and 8 were the same as Conditions 2-6: responses in the first 3 s continued to have no scheduled consequences, while responses after these initial 3 s could be followed by food. Only the key that delivered the response-contingent stimulus was red; the other key remained white. In both Conditions 7 and 8 the probability that the next food would be on the same alternative as the last red keylight was .1. Red keylights were therefore unlikely to acquire conditioned value via pairing — approximately 90% of foods followed a (peck to a) white keylight, not a red one. Additionally, increased control by the red keylights in Conditions 7 and 8 would manifest as greater preference to the not-just-productive alternative, and thus could not result from any increases in hedonic value brought about via the (relatively infrequent) instances in which red keylights preceded food.

Extending the red keylights may lead to increases in sensitivity to the local obtained food ratio first, by eliminating the need to actively remember the location of the most recent red keylight (reducing working memory load), and second by allowing for greater discrimination of the relationship between red keylights and foods. In addition to reference and working memory, Honig (1978) also identified what he termed *associative memory,* the memory required for, and involved in, associating two temporally distant events. Associative memory is claimed to be involved in both operant processes such as delayed reinforcement, and Pavlovian processes such as long-delay conditioning and taste-aversion learning. In Experiment 2a, the food ratios obtained at relatively long times from the last red keylight might simply have never become associated with that temporally distant red keylight.

Although response-reinforcer contingencies have been learned with delays of up to 30 or even 60 s (Critchfield & Lattal, 1993; Lattal & Gleeson, 1990; Williams & Lattal, 1999), and even 24 hours (Ferster & Hammer, 1965), other evidence suggests that the contingency between response and reinforcer is less well discriminated when the intervening delay is long. For example, response rate is generally an inverse function of response-reinforcer delay (Odum, Ward, Barnes, & Burke, 2006;

Richards, 1981), and responding is more likely to be maintained on additional, non-functional manipulanda when there is a delay between reinforcement and responses on the functional manipulandum (Escobar & Bruner, 2007; Keely, Feola, & Lattal, 2007). Thus, while it would be possible for the relationship between response-contingent red keylights and foods to be learned, even at relatively long delays, this relationship would be less likely learned than if foods closely followed the red keylights. Further, there is no guarantee that the relation between red keylight and obtained food ratio learned at times close to the most recent red keylight would generalize to other, longer times. White and Cooney (1996) separately manipulated the red: green reinforcer ratio at two delays (0.1 s and 4 s) in a delayed matching to sample (DMTS) task. They found that varying the reinforcer ratio at the long delay had no effect on the response ratio at the short delay and vice versa (see also Sargisson & White, 2001), implying independence in stimulus control by recent, and temporally distant stimuli.

Conditions 9 and 10 were direct replications of Conditions 2 and 3 respectively. If extending the stimuli had no effect other than to remove the need to remember the location of the last stimulus, the results of Conditions 2 and 3 should be exactly replicated in Conditions 9 and 10. Thus, low measures of stimulus control at long times since the last red keylight would again be expected. On the other hand, experience of the red keylight extending until the next food in Conditions 7 and 8 may facilitate learning that the red keylight is relevant to the local food ratio even at relatively long temporal distances from its initial presentation. This learning of the relevance of the red keylight to the food ratio at all times after its delivery could persist beyond the stimulus extension conditions and thus cause more extreme preference in Conditions 9 and 10 than in Conditions 2 and 3.

4.3.1 Method

Subjects and Apparatus

The subjects and apparatus from Experiment 2a were again used.

Procedure

In Conditions 7 and 8 the red keylight stimuli were extended beyond the 3 s of Experiment 2a. In all other respects, the basic procedure in Conditions 7 and 8 was unchanged from the procedure of Experiment 2a: Response-contingent events were delivered according to an overall VI 27-s schedule and foods and red keylights strictly alternated. The left versus right location of each food was determined according to probabilities unique to that condition.

The first 3 s of a red keylight delivery were the same in Conditions 7 and 8 as in Experiment 2a: that is, no other keylight in the chamber was illuminated, and responses had no scheduled consequences. After these 3 s, the switching key was turned on and responses to it and the red side-key were again effective. Pecks to the red keylight could be followed by food (although this was scheduled to only happen for 10% of the total foods). The switching key continued to operate in the same way as in Experiment 2a. Whenever the side key that had provided the last red keylight came on, it was red. The other side-key was lit white whenever it was on.

The overall left:right red keylight ratio was 1:1 in Condition 7 and 9:1 in Condition 8. In both conditions, the local probability of a same-alternative food after a red keylight was .1 meaning that in Condition 7 the overall food ratio was 1:1 while in Condition 8 this ratio was 1:9. Conditions 7 and 8 were each conducted for 49 sessions.

Conditions 9 and 10 were replications of Conditions 2 and 3 respectively: the red keylights again only lasted 3 s and each condition lasted 65 sessions. In both conditions, the overall food and red keylight ratios were 1:1. The local probability of a same-alternative food after a red keylight was .9 in Condition 9 and was .1 in Condition 10. Each session of each condition lasted 60 reinforcers or 60 minutes, whichever occurred first.

4.3.2 Results

Although fewer sessions were conducted in Conditions 7 and 8, behaviour stabilized more quickly (see Appendix D) so data could still be taken for analysis from the final 20 sessions. Figure 4.7 presents the group mean log (left/right) response ratio in successive 2-s time bins after each of the four response-contingent events. These preference pulses were created in the same way as were the preference pulses in Experiment 2a. The individual subject preference pulses for all conditions of Experiment 2b are presented in Appendix E (Figures E1-E6)

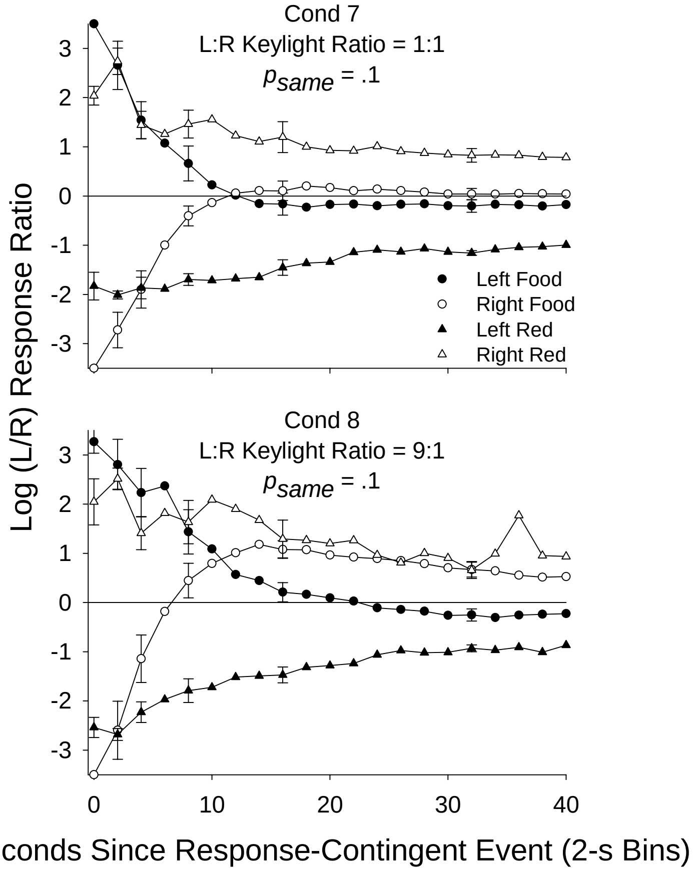


Figure 4.7. Group mean log (L/R) response ratio in each successive 2-s time bin after each of the 4 response-contingent events in Conditions 7 and 8 (stimuli extended to last until the next food). Error bars are plotted at representative data points and depict 1 standard error.

In both Conditions 7 and 8, preference immediately after a response-contingent red keylight was strongly towards the not-just-productive alternative and it remained that way for more than 40 s. These preference pulses differed noticeably from those obtained in Experiment 2a and indicate that responding was roughly in accordance with the local food ratio for the entire period after the initial 3-s red keylight. Preference pulses for Conditions 9 and 10 (in which the red keylights were again shortened to 3 s) are presented in Figure 4.8 along with the preference pulses from Condition 2 and 3 (which Conditions 9 and 10 replicated) for comparison.

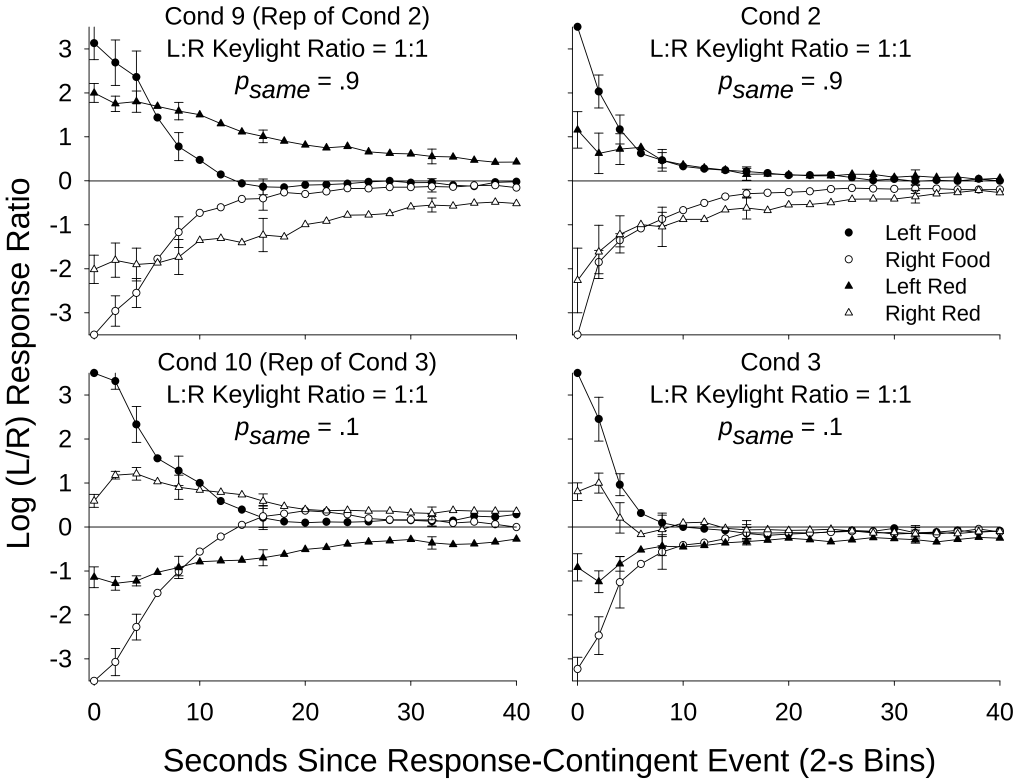


Figure 4.8. Group mean log (left/right) response ratio as a function of successive 2-s time bins after each of the 4 response-contingent events in Conditions 9 and 10 (left panels). Preference pulses from Conditions 2 and 3 (Experiment 2a) are reprinted here to aid in comparison.

Although procedurally, Condition 9 was identical to Condition 2 and Condition 10 was identical to Condition 3, the results were not replicated. Unlike Condition 2, preference in Condition 9 remained toward the just-productive alternative up to 40 s after a red keylight. Preference in Condition 10 remained toward the not-just-productive alternative for much longer than in Condition 3. Thus, the local behaviour ratio in the post-red keylight period was generally closer to the local arranged food ratio in that period in the two conditions which followed the stimulus extension conditions. What of the local post-red keylight obtained food ratios? Figure 4.9 presents the log (left/right) response and response-contingent event ratios for all conditions of Experiment 2b (the individual subject plots are presented in Appendix Figures E7-E12).

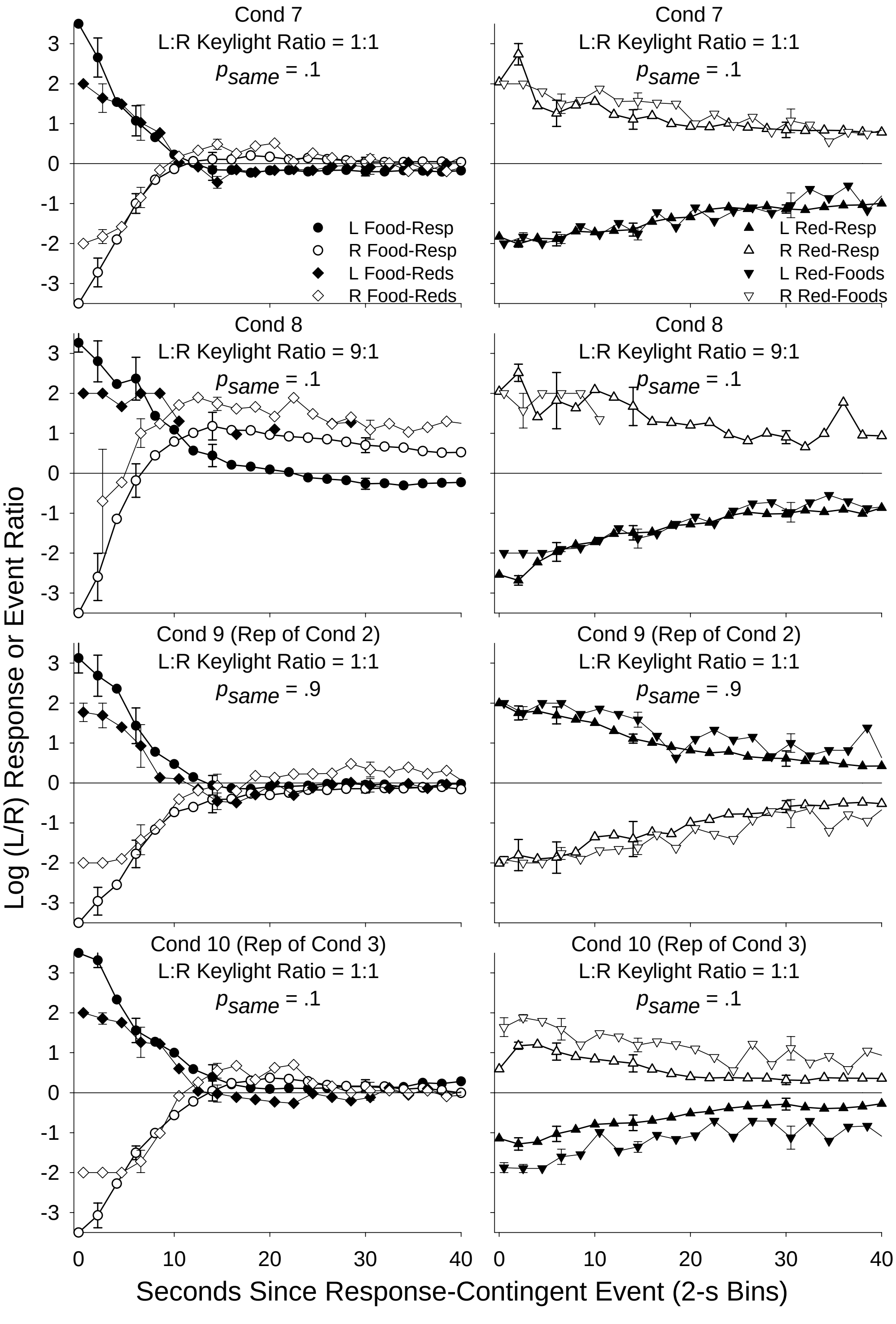


Figure 4.9. Group mean log (L/R) response and response-contingent event ratios in all conditions of Experiment 2b. The left panel shows the response and contingent event pulses after a left or right food. The right panel shows these pulses after a left or right red keylight. Error bars depict one standard error and are plotted at representative data points that increment in $\log_2$ units. The contingent event pulses are shifted to the right of the response pulses by 0.5 s on the x-axis and are connected with a thinner line in order to aid comparison of the plots.

Extending the red keylights (Conditions 7 & 8) led to a closer approximation of the local response ratio to the local obtained food ratio in the post-stimulus period. This closer approximation carried over into conditions when the stimuli were again shortened (Conditions 9 & 10). The local response ratios appear more similar to the local obtained food ratios in Conditions 9 and 10 (Figure 4.9) than they were in Conditions 2 and 3 (Figure 4.2). Sensitivity of the local behaviour ratio to the local obtained food ratio after each response-contingent event was calculated for Conditions 7 and 8 and is presented in Figure 4.10.

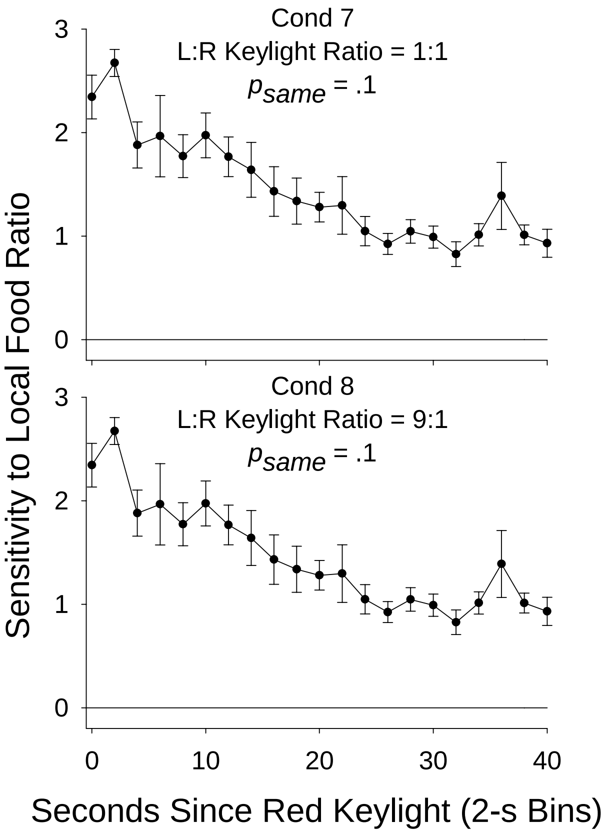


Figure 4.10. Group mean sensitivity of the local behaviour ratio in each 2-s time bin after a response-contingent red keylight to the obtained food ratio in that period in Conditions 7 and 8 (stimulus extension conditions). Error bars depict 1 standard error.

Sensitivity of the local behaviour ratio to the local obtained food ratio was higher for a longer time in Conditions 7 and 8 than in Experiment 2a (Figure 4.4). This was expected given the greater similarity of the local response and obtained food ratios in the latter conditions (Figure 4.9). Estimates of sensitivity of the local behaviour ratio after a red keylight to the local obtained food ratio in that period in Conditions 9 and 10 are presented in Figure 4.11 along with the (reprinted) plots from Experiment 2a.

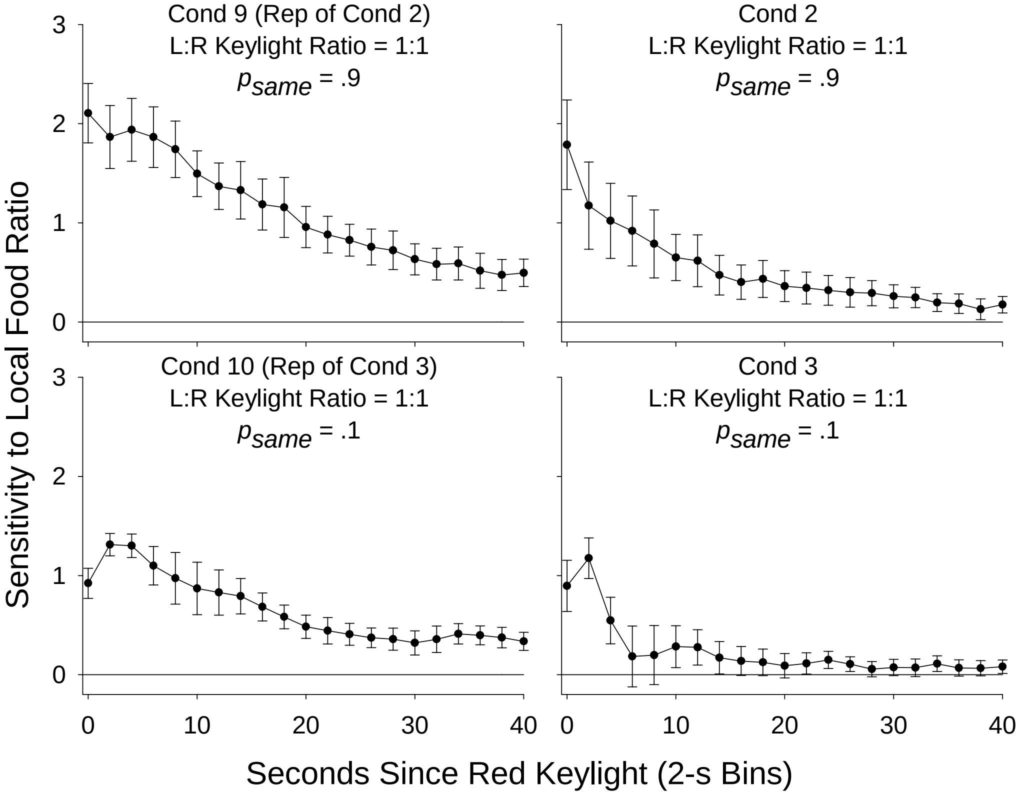


Figure 4.11. Group mean sensitivity of the local behaviour ratio in each 2-s time bin after a response-contingent red keylight to the obtained food ratio in that period in Conditions 9 and 10 (Experiment 2b; left panels) and Conditions 2 and 3 (Experiment 2a; right panels). Error bars depict 1 standard deviation.

Choice was more sensitive to the local food ratio after a red keylight in Conditions 9 and 10 (Experiment 2b) than in Conditions 2 and 3 (Experiment 2a). In particular, sensitivity to the local obtained food ratio was higher at times further from the red keylight in the latter conditions. Consistent with Experiment 2a, sensitivity to the local food ratio was higher when the local probability of a same-alternative reinforcer after a red keylight was .9 (Condition 9) than when it was .1 (Condition 10).

4.3.3 Discussion

Condition 9 was procedurally identical to Condition 2 and Condition 10 was procedurally identical to Condition 3. The different results thus could not be due to different procedures. More likely, they were due to the intervening history between the original and the replication conditions. Choice immediately after a red keylight remained at preference for the alternative more likely to produce the next food for a longer period in Conditions 9 and 10 than in Conditions 2 and 3. Preference was also further towards the alternative more likely to deliver the next food in Conditions 7 and 8. The experience in Conditions 7 and 8 is thus very likely responsible for the failure of Conditions 9 and 10 to replicate the results of Conditions 2 and 3.

In Conditions 7 and 8, preference to the alternative more likely to provide the next food was manifested as greater responding to the not-just-productive alternative. In Condition 9, this preference was manifested as greater responding to the just-productive alternative. Thus, a simple rule such as “avoid the red keylight” is not likely to have been learned in Conditions 7 and 8 and transferred into Condition 9. The mean interval between a red keylight and food was calculated for each condition in both Experiments 2a and 2b and these are presented in Table 4.2. In Conditions 2 through 6 (Experiment 2a) and Conditions 9 and 10 (Experiment 2b), the interval was simply the time from the end of a 3-s red keylight and the start of the next food. In

Conditions 7 and 8 the interval was 0 if a food was delivered for a peck to the alternative that delivered the last red keylight, and was the interval from the third changeover peck (which turned the previously pecked key off) to food if that food was delivered for a peck to the not-just-productive alternative.

Table 4.2

Mean (and standard deviation) red keylight-food interval (in seconds) for each individual subject and the group in each condition of Experiments 2a and 2b.

Condition	Pigeon 11	Pigeon 12	Pigeon 13	Pigeon 14	Pigeon 15	Pigeon 16	Group Mean
2	31.10	29.13	29.96	31.02	30.38	30.73	30.39
	(4.69)	(4.87)	(4.89)	(4.96)	(5.14)	(4.87)	(4.90)
3	31.35	31.74	31.52	31.90	33.62	35.77	32.65
	(4.64)	(4.57)	(4.25)	(4.80)	(4.71)	(4.32)	(4.55)
4	29.86	32.68	30.62	31.61	32.51	32.36	31.61
	(4.36)	(4.97)	(4.94)	(4.55)	(5.29)	(5.04)	(4.86)
5	32.10	32.58	30.42	31.44	33.16	32.74	32.07
	(5.42)	(4.58)	(4.52)	(4.41)	(4.56)	(4.58)	(4.68)
6	29.50	31.15	30.55	32.55	30.90	32.11	31.13
	(4.36)	(4.93)	(4.41)	(4.55)	(4.38)	(4.66)	(4.55)
7	25.04	25.41	26.77	23.70	23.64	27.58	25.36
	(5.06)	(5.10)	(5.32)	(4.43)	(4.70)	(5.32)	(4.99)
8	25.71	26.23	24.52	26.21	24.29	24.63	25.30
	(5.40)	(5.18)	(5.07)	(4.79)	(4.56)	(4.99)	(5.00)
9	30.36	31.74	30.48	31.11	30.24	29.32	30.54
	(4.72)	(5.09)	(4.67)	(4.65)	(4.72)	(5.08)	(4.82)
10	30.08	30.33	33.22	31.12	29.40	31.87	31.01
	(4.63)	(4.74)	(4.83)	(4.86)	(4.75)	(4.89)	(4.78)

For each individual subject, the mean stimulus-food interval in Conditions 7 and 8 was lower than the mean interval in any of either the preceding or the following conditions. The differences however were all rather small (Mean 6.02 s) and only slightly larger than the standard deviations (Mean 4.79 s). This finding eliminates the pairing hypothesis of conditional reinforcement as the cause of the increased preference in Conditions 7 and 8: the average interval between a red keylight and food was still at least 23 s even in these conditions. The mean stimulus-food interval again increased in Conditions 9 and 10 to values comparable to those in Experiment 2a. Yet, the preference pulses in these latter conditions were still more extreme relative to the preference pulses in the earlier conditions (Figure 4.8). These findings also make it unlikely that extending the stimuli reduced working memory load. More likely, extending the stimuli in Conditions 7 and 8 facilitated learning that the red keylights signalled the local food ratio even at relatively long times since that red keylight. This influence of the red keylight on behaviour at relatively long temporal distances from its initial illumination persisted even when the red keylight duration was again shortened in Conditions 9 and 10. The stimuli in Experiments 2a and 2b were equivalent in all other respects: they signalled the same average reduction in time to food, they signalled the same local food ratio and the information provided by the stimuli was equally relevant.

4.4 Experiment 2c

The results of Experiment 2b account for the relatively small preference pulses in conditions of Experiment 2a where the keylights signalled the likely location of future food (Conditions 2-4). They did not address the relatively extreme preference pulses that followed stimuli which signalled nothing about the likely location of future food (Conditions 5 & 6). These relatively large preference pulses could be due to the fact that although the red keylights signalled nothing about the likely location of the next food in Conditions 5 and 6, they did signal that food was now temporally closer than it had been prior to that red keylight. This signalled reduction in waiting time to food may have endowed the stimuli with conditionally reinforcing properties (Squires & Fantino, 1971).

Alternatively, the relatively extreme local preference to the just-productive alternative in Conditions 5 and 6 may be at least partly attributable to carry-over from the previous condition. There was a high probability that the just-productive alternative would produce the next food in Condition 4. Local preference is apparently highly sensitive to previously experienced, as well as current, contingencies. Krägeloh and Davison (2003) reported flat preference pulses from initially naïve animals working in a frequently-changing concurrent schedule with no changeover requirement. Introducing a changeover delay led to preference pulses, and these persisted even when the changeover requirements were later removed. In Experiment 2a, local preference may have been influenced by the local food ratios signalled by the red keylight in earlier conditions. Experience in earlier conditions may have created some initial tendency to stay at the just-productive alternative in Condition 5. This tendency became entrenched by the interaction between the local response and obtained food ratios throughout Condition 5 (Figure 4.3). Perhaps if Condition 6 (red keylights not informative about the local food ratio) was conducted first, choice after a red keylight would have been indifferent. For this reason, a second stimulus with no history of signalling the next food's location (green keylight) was introduced in Experiment 2c.

In Conditions 11-14 a second stimulus was added. In Conditions 11 and 13 (which were procedurally equivalent to one another), red keylights signalled that food was very likely ($p = .9$) to appear on the just-productive alternative (as they did in Conditions 2 & 9). Unlike Conditions 2 and 9, a random half of all response-contingent keylights were green. The probability that a food would be on the same alternative as the last green keylight was .5. Red keylights were informative in both the temporal sense (a food was now temporally closer than it had been prior to the red keylight) *and* the response-requirement sense (red keylights also signalled that the next food was very likely to be on the just-productive alternative). The green keylights were only informative in the temporal sense. Additionally, these green keylights had no history of being informative on the likely location of future foods. In Condition 12 the probability of a same-alternative reinforcer after a green keylight remained .5, but the probability of a same-alternative reinforcer after a red keylight was changed to .1 (as in Conditions 3 and 10). In Condition 14, the green keylights were, for the first time, made informative about the likely location of the next food. In this condition, the probability of a same-alternative food was .9 after a red keylight and was .1 after a green keylight.

4.4.1 Method

Subjects and Apparatus

The subjects and apparatus from Experiments 2a and 2b were again used.

Procedure

As in previous conditions, response-contingent changes in keylight colour strictly alternated with response-contingent foods. A single overall VI 27-s schedule continued to arrange both the keylights and the foods, and all response-contingent events continued to be 3-s long. However, unlike previous conditions, the response-contingent changes in keylight colour were divided into red keylights, which these pigeons had extensive experience with, and green keylights, which these animals were seeing for the first time. In Conditions 11 and 13 (which were procedurally identical), as well as Condition 14, a red keylight signalled that the next food would be on the same alternative as the last red keylight with a probability of .9. In Condition 12, a red keylight signalled a probability of .1 that the just-productive alternative would provide the next food. In all conditions of Experiment 2c, a random half of all keylights were red. The other half were green and signalled that the next food would be on the just-productive alternative with a probability of .5 in Conditions 11 to 13, and that the next food would be on the just-productive alternative with a probability of .1 in Condition 14. As in Experiments 2a and 2b, each session lasted 60 seconds or 60 food-deliveries (whichever occurred first) and each session lasted 65 sessions.

4.4.2 Results

The justification for continuing to analyse the last 20 sessions of each condition is presented in Appendix F. The log (left/right) response ratio after each response-contingent event in Conditions 11 to 14 were created in the same way as in previous conditions. These are presented for the individual subjects in Appendix Figures G1-G6 and for the group in Figure 4.12.

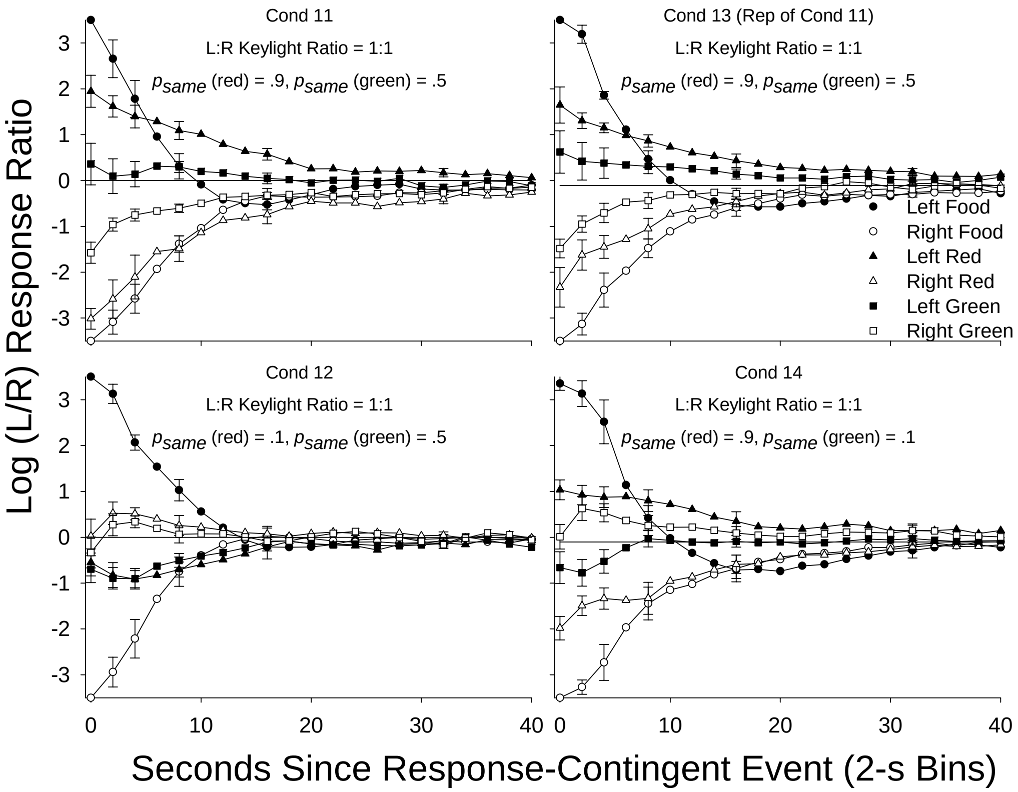
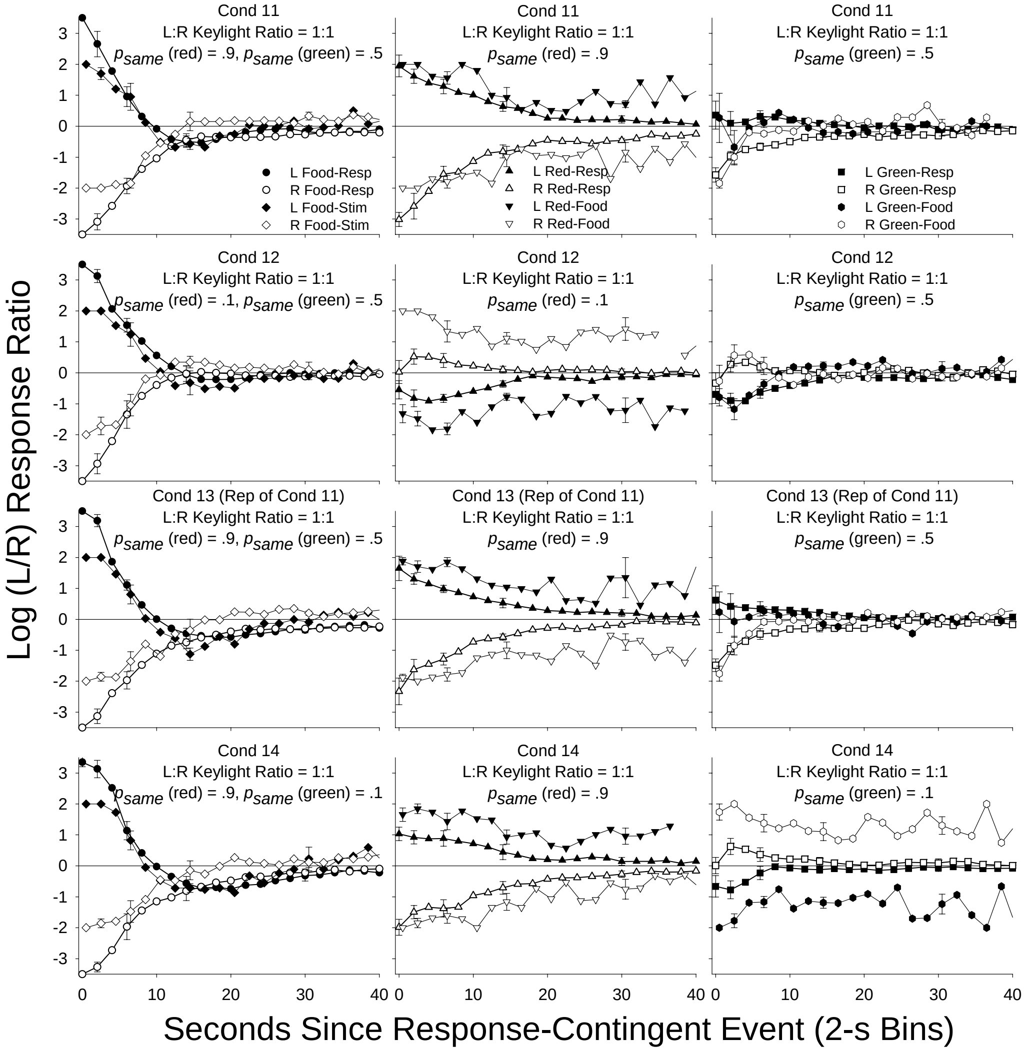


Figure 4.12. Group mean log (L/R) response ratio in each successive 2-s bin after each of the 6 response-contingent events in Conditions 11-14. Error bars are plotted at representative data points and depict 1 standard error.

In Condition 11, as in earlier conditions with a .9 probability of a same-alternative food after a red keylight, preference after a red keylight was towards the just-productive alternative. Preference after a green keylight in Condition 11 was also towards the just-productive alternative, but was noticeably less extreme than preference after a red keylight. Condition 13 closely replicated Condition 11: preference was further towards the just-productive alternative after a red than after a green keylight. There was perhaps some small tendency for preference to be less extreme in the later condition, but the same qualitative patterns were present. In Condition 12 the probability of a same-alternative food after a red keylight was changed to .1 (from .9 in Condition 11), while the probability of a same-alternative food after a green keylight was unchanged (p =.5). Local preference after *both* red and green keylights was towards the not-just-productive alternative. This suggests that there may have been some blending of the local reinforcer ratios signalled by the red and green keylights. In Condition 14, red keylights signalled a .9 probability that the next food would be on the same alternative, and green keylights signalled a .1 probability. Preference after each of these events was clearly a function of these local contingencies. Preference after a red keylight was towards the just-productive alternative while preference after a green keylight was towards the not-just-productive alternative. The interaction between the local response ratios after red and green keylights (especially in Condition 12) suggests that the local obtained food ratios may have differed from the local arranged food ratios. Figure 4.13 presents the preference and response-contingent event pulses in Conditions 11-14. There are 3 plots per condition in Figure 4.13 because of the additional green keylights. Red and green keylights were not differentiated in calculating the obtained log keylight ratio following left and right foods (left panels). The individual subject plots are presented in Appendix Figures G7-G12.



4.13 Group mean log (L/R) response and response-contingent event ratios in Conditions 11-14. The left panels show behaviour and keylights after foods, the middle panels show behaviour and foods after red keylights, and the right panels show behaviour and foods after green keylights. Error bars depict one standard error and are plotted at representative data points that increment in $\log_2$ units. The contingent event pulses are shifted to the right of the response pulses by 0.5 seconds on the x-axis and are connected with a thinner line in order to aid comparison of the plots.

The local response ratios after left and right, foods and red and green keylights in Condition 11 were similar to the local obtained response-contingent event ratios in those periods. Immediately after a red keylight, the local behaviour and food ratios were both towards the just-productive alternative before approaching (though not quite reaching) indifference. The local behaviour ratios approached indifference at a somewhat faster rate than the local obtained food ratios. After a green keylight, the local behaviour and obtained food ratios were generally similar to one another and were both towards the just-productive alternative (although to a lesser degree than after a red keylight). All of these trends were generally replicated in Condition 13, although preference after a red keylight was somewhat less extreme in the later condition.

The preference and response-contingent event pulses in Condition 12 were more dissimilar to one another. In Condition 12, the local obtained food ratio after a red keylight was strongly towards the not-just-productive alternative, while the local behaviour ratio in that period was only somewhat towards the not-just-productive alternative. This was also true in earlier conditions with a .1 probability of a same-alternative reinforcer. The local response and food ratios were more similar after green keylights in Condition 12. Although the arranged local reinforcer ratio in that period was 1:1, there was a clear trend for both preference and the local obtained food ratio to be somewhat towards the not-just-productive alternative. Local preference after a red keylight in Condition 14 (when $p_{same} = .9$) was again similar to the local obtained food ratio. After a green keylight however ($p_{same} = .1$), preference was much less extreme than the local obtained food ratio. The local behaviour and food ratios after green keylights in Condition 14 were very similar to those after red keylights in Condition 12 (which also signalled a .1 probability of a same-alternative food).

4.4.3 Discussion

In Experiment 2c a new stimulus was introduced which was uninformative about the likely location of future food and had no history of being informative in this regard. These green keylights did however signal that food was temporally closer.

Local preference after a green keylight in Conditions 11 and 13 was less extreme than was local preference after a red keylight in Condition 6, despite the fact that these stimuli all signalled the same local probability of a same-alternative reinforcer (p_{same} = .5). This may be because the green keylights in Conditions 11 and 13 had never before signalled anything about the likely location of future foods. The red keylights in Condition 6 on the other hand had preceding histories of signalling food's likely location.

When the red keylights signalled a low probability (p = .1) of a same-alternative food (Condition 12), not only was preference after a red keylight towards the not-just-productive alternative, but so too was preference after a green keylight, even though green keylights signalled a .5 probability of a same-alternative reinforcer. In no other condition with a 1:1 local reinforcer ratio was there any indication of local preference tending towards the not-just-productive alternative. Preference after a green keylight may have been towards the not-just-productive alternative in Condition 12 because of generalization across red and green keylights. The .1 probability of a same-alternative reinforcer signalled by a red keylight, and the .5 probability signalled by a green keylight may have been combined into something approximating a .3 probability of a same-alternative reinforcer after any colour keylight.

Generalization was not wholly responsible for local preference after a green keylight however. Figure 4.13 demonstrated that the local obtained food ratio after a green keylight was also somewhat towards the not-just-productive alternative. The slight tendency for the animals to switch after a green keylight in Condition 12 drove the local obtained food ratio towards the not-just-productive alternative. This, once again indicates a dynamic system in which an initial tendency to switch was followed by a local obtained food ratio that deviated from that arranged (1:1). This local obtained food ratio then further reinforced the tendency to switch.

Generalization of the reinforcer ratio signalled by red and green keylights also accounts for the fact that preference after a red keylight was less extreme in Condition 12 than in any other condition with a .1 signalled probability of a same-alternative reinforcer. Preference after a red keylight was also a function of the (combined) .3 probability of a same-alternative reinforcer. The undifferentiated control by the keylight colours was apparently restricted to Condition 12 however. In Condition 13, which was a replication of Condition 11, preference was again differentially controlled by the red and green stimuli. Also, in Condition 14, when red stimuli signalled a .9 probability of a same-alternative reinforcer and green stimuli signalled a .1 probability, preference was apparently a function of these two distinct local probabilities of a same-alternative food. Local preference was not indifferent, as would be the case if the .9 and .1 probabilities of a same alternative reinforcer were averaged. Thus, distinct discriminative control by two stimuli seems to disappear only when one stimulus is uninformative about the likely location of the next food (p_{same} = .5), and the other stimulus signals a low probability that this next food will appear on the just-productive alternative (p_{same} = .1).

4.5 General Discussion

The results of Experiment 2 confirm that choice after response-contingent stimuli in Experiment 1 differed from choice after such stimuli in Davison and Baum's (2006; 2010) experiments because of differences in what the stimuli signalled about forthcoming reinforcement. While the paired stimuli in Experiment 1 signalled the only period in which the forthcoming local food rate was greater than zero, the stimuli in Davison and Baum's experiments uniquely and non-redundantly signalled what response (left vs. right) was more likely to produce the next food. In Experiment 2, the response-contingent keylight stimuli always signalled that food was temporally closer, and in some conditions also signalled the likely location of the next food.

Unlike Experiment 1 (and as in Davison & Baum's, 2006; 2010 experiments) an effect of signalling the likely location of the next food was found in Experiment 2. When the probability of a same-alternative reinforcer was either .9 or .1, preference immediately after a response-contingent keylight was towards the alternative more likely to provide the next food. This control by the signalled local food ratio was rather short-lived in Experiment 2a: preference stabilized at a level reflective of the overall food ratio rather than the food ratio particular to the period after the last red keylight.

When the stimuli were extended in Experiment 2b, preference after a red keylight remained towards the alternative more likely to provide the next food for longer. Extending the stimuli only marginally decreased the average interval between a red keylight and food (Table 4.2) and sensitivity to the local contingencies remained high even when the stimuli were again shortened (Figure 4.11), indicating that the observed increase in sensitivity to the local food ratio was not due to the extended presence of the red keylights per se. Instead, extending the stimuli facilitated learning about the relevance of the stimuli to the local food ratio at relatively long temporal distances from the last red keylight. Experiment 2b indicated that behaviour could be controlled by temporally distant stimuli, if only the contingencies are adequately discriminated.

As Conditions 9 and 10 were procedurally identical to Conditions 2 and 3 respectively, the differences between those conditions could be considered a failure to replicate. Sidman (1960), however, argued that a failure to obtain within-subject replication does not necessarily indicate a lack of control if accompanied by inter-subject replication. Such situations indicate that behaviour is changing across time independently of the particular manipulations being investigated. The effects of the manipulations should be evaluated against this changing baseline. In this case, preference was more extreme after the extension conditions (Conditions 9 & 10) than before it (Conditions 2 & 3), whether the red keylight signalled a high (.9) or a low (.1) probability of a same-alternative food. The effects of signalling the local food ratio should thus be evaluated relative to this increasingly more extreme preference. In both Condition 2 and its replication (Condition 9), preference immediately after a red keylight was towards the just-productive alternative. Preference in Condition 3 and its replication (Condition 10) on the other hand, was towards the not-just productive alternative. Thus, the main finding of Experiment 2a was replicated: preference after a red keylight was towards the alternative signalled by that red keylight as more likely to provide the next reinforcer.

Experiments 2a and 2b also indicated that, despite the long-running conditions and extended periods allowed for behaviour to stabilise (45 sessions), local preference was affected by historical contingencies. When the stimuli were extended, behaviour at relatively long times since the start of the keylight came under the control of that keylight’s location. This continued even when the keylight stimuli were again shortened in Conditions 9 and 10.

History effects were not limited to Conditions 9 and 10 being more extreme than Conditions 2 and 3. Red keylights signalled a 1:1 local food ratio in Conditions 5 and 6. Despite this, preference immediately after a red keylight in those conditions was reliably and strongly towards the just-productive alternative. This may have been because of the immediately prior condition where red keylights signalled a .9 local probability of a same-alternative reinforcer (Condition 4). Weiner (1964; 1969) found that whether human subjects were previously exposed to a differential reinforcement of low rates (DRL) schedule of reinforcement, or a fixed ratio (FR) schedule, determined their behaviour in a subsequent FI schedule — those previously exposed to a DRL responded at low rates while those previously exposed to an FR responded at high rates. These results were replicated to some degree within subjects (Weiner, 1969, Experiment 3) as well as with pigeons (Freeman & Lattal, 1992) and rats (Johnson, Bickel, Higgins, & Morris, 1991). According to Weiner (1969), behavioural patterns were maintained from one condition to the next because the new contingencies (reinforcement according to an FI schedule) reinforced this persistence: response rates that are either too fast or too slow impose no penalties, in terms of delaying reinforcers, in FI schedules. DRL effects do not persist into FR schedules and FR effects do not persist into DRL schedules precisely because, in those schedules, reinforcer rate is directly related to response rate.

Adherence to historical contingencies can persist even when the new contingencies do directly penalize such adherence. Weiner (1970) noted that responding previously reinforced according to an FR schedule remained at a high rate even when a response cost was added to the new FI schedule. Something similar may have occurred in Conditions 5 and 6. Previous exposure to Condition 4, where the probability of a same-alternative reinforcer after a red keylight was .9, established a pattern whereby the local response ratio after a red keylight was strongly towards the just-productive alternative. When the local reinforcer ratio after a red keylight was changed in Condition 5, there may have been a small cost imposed on the continuing extreme preference immediately after a red keylight (e.g., reinforcers on the not-just-productive alternative may have been somewhat delayed). This small cost however was insufficient to counteract the previous history in which responses to the just-productive alternative were more likely reinforced. Arranging a .5 local probability of a same-alternative reinforcer may be a *weak contingency* (Davison, 1998) which does not strongly drive behaviour and readily reflects prior history, in this case, the immediately prior .9 local probability of a same-alternative reinforcer. An established preference or pattern of behaviour could be changed by current contingencies if those new contingencies were strong enough. This was the case in Condition 3 (which arranged a .1 local probability of a same-alternative reinforcer and followed a condition where this probability was .9) but not Condition 5 (where this probability was .5).

The relationship between local response and food ratios was not unidirectional, with the experimenter-defined local food ratio controlling choice. Rather, the local obtained response and food ratios existed within a dynamic system where the local behaviour ratio influenced the local food ratio and vice-versa. Thus, although the arranged local food ratio throughout the post-red keylight period was 1:1 in Conditions 5 and 6, the relatively extreme (although transient) preference to the just-productive alternative drove the local food ratio in the first 10 s after a red keylight away from this 1:1 ratio. This relatively extreme food ratio then maintained a relatively extreme local response ratio. Davison (1998) noted the difficulty in predicting behaviour in such systems where the obtained reinforcers are dependent on emitted behaviour. Thus, although generalization from the prior condition may have been *initially* responsible for the preference pulse to the just-productive alternative in conditions arranging a .5 probability of a same-alternative reinforcer, generalization was not wholly responsible — the continuing extreme local obtained food ratios maintained the extreme local preference.

Because historical contingencies exerted relatively strong effects, new stimuli that had no history of signalling the likely location of future food were introduced in Experiment 2c. These green keylight stimuli only signalled that food was temporally closer, and were presented along with red keylights which continued to signal both that food was temporally closer and the likely location of that food. When these green keylights were presented along with red keylights that signalled a .9 probability of a same-alternative food (Conditions 11 & 13), preference for the alternative that produced the last green keylight was both small and short-lived. This confirmed that the previously obtained, apparently large, effects of signalling only that food was closer in time (Conditions 5 and 6 of Experiment 2a) were more a reflection of historical contingencies where the location of the next food was signalled, than of current contingencies where the only thing signalled was that food was temporally closer.

When the local probability of a same-alternative reinforcer after a red keylight was changed to .1 (the probability of a same-alternative reinforcer after a green keylight remained .5), preference after both red *and* green keylights shifted towards the not-just-productive alternative. This suggests some generalization of the local probability of a same-alternative reinforcer across keylight colours. Although the local probability of a same-alternative reinforcer was .1 after a red keylight and was .5 after a green keylight, preference after any keylight, red or green, appeared to be controlled by a (.1+.5)/2 = .3 probability of a same-alternative reinforcer. This explains not only why preference after a green keylight was towards the not-just-productive alternative, but also why preference after a red keylight was in the direction predicted by the local food ratio, but was attenuated relative to prior conditions arranging the same local probability of a same-alternative reinforcer.

Just as generalization across conditions influenced the local post-keylight preference and obtained food ratios in Conditions 5 and 6, generalization across keylight colours influenced both the local preference and obtained food ratios in Condition 12. Krägeloh, Elliffe and Davison (2006) also found that the reinforcer ratio signalled by one stimulus generalized to behaviour in the presence of other, similar stimuli. They arranged a frequently-changing procedure where each component was signalled by red and yellow keylight flashes (Krägeloh & Davison, 2003). In some conditions, 5 of the 7 components arranged a 1:1 food ratio while the other two components arranged food ratios of 27:1 and 1:27. Krägeloh et al. reported that, despite long-running conditions, responding generalized from conditions with a 27:1 or 1:27 food ratio to conditions with a 1:1 food ratio and similar discriminative stimuli. Both the preference trees and the preference pulses in the 1:1 components were ordinally related to the similarity of the component discriminative stimulus to the discriminative stimuli in the 27:1 and 1:27 components. A local food ratio of 1:1 is indeed a weak contingency (Davison, 1998), readily reflecting both historical contingencies and current contingencies signalled by other stimuli.

What can the results of Experiment 2 say about conditional reinforcement and the role of putative conditional reinforcers signalling the time and likely location of future primary reinforcers? First note that the red keylights were only paired with food in 2 of the 14 conditions (Conditions 7 and 8) and even in those conditions red keylight-food pairings were rare (only 10% of foods in those conditions followed a red keylight). Despite this, a red keylight produced a transient increase in preference to the just-productive alternative whenever it did not signal that food was more likely on the other alternative. Any account which requires that a stimulus be temporally contiguous with primary reinforcement in order for this stimulus to engender apparent reinforcer effects cannot account for these results (Williams, 1991b, 1994).

Might the keylights have acquired conditioned value through some means other than temporal contiguity with food? The notion that stimuli which signal a reduction in time to primary reinforcement themselves acquire value as reinforcers has been formalized a number of times (Killeen, 1982; Mazur, 2001; Squires & Fantino, 1971). Although the stimuli always signalled a reduction in time to primary reinforcement, local preference after a stimulus was not always indicative of that stimulus having any acquired value: when the keylights signalled that food was likely to come next for a response to the not-just-productive alternative, there was a relative *decrease* in responses to the alternative that preceded the stimulus. Such a relative decrease in responses to the alternative that produced the stimulus is not consistent with any account of the stimulus acting as a reinforcer. Additionally, Experiment 2c demonstrated that preference for an alternative which produces stimuli that only signal food's imminence is better attributed to the influence of historical contingencies which, along with the changeover requirement, pushed preference to the just-productive alternative. This preference was then maintained via a dynamical relationship between the local response and obtained food ratios.

The red and green keylights may thus be best understood as discriminative stimuli which signalled the likely location of the next food. Taken together, Davison and Baum's (2006; 2010) experiments, Experiment 1 and Experiment 2 offer a comprehensive conceptualization of conditional reinforcement in operant procedures. Although biologically neutral stimuli may acquire hedonic properties via their relationship with intrinsically appetitive stimuli (these experiments are silent on this), it is not these properties that are responsible for the behaviour before or after the stimuli. Rather, the discriminative or informative properties are responsible for the behavioural effects of the stimuli. If the stimulus signals that the preceding behaviour is highly likely to produce food (or it did so historically), then it will appear to be a conditional reinforcer — an increase in preference to the alternative that produced it will follow. This is insufficient to claim definitively that the stimulus is indeed a reinforcer however. Just as behaviour can be driven by the stimulus towards the immediately preceding response, it can also be driven away from this response.

This understanding of conditional reinforcer effects does not explicitly claim that the conditional reinforcer has no (acquired) appetitive properties; only that these properties are not responsible for the behavioural effects seen in this and other operant procedures. Indeed, compelling evidence exists that motivational properties can be transferred from one stimulus to another. For example, rats demonstrate preference for a location paired with morphine but an aversion to morphine-paired flavours. If the flavour is first paired with morphine, then with a location (i.e., second-order conditioning), rats demonstrate an aversion to the location. Similarly, if the location is first paired with morphine, then with flavour, a preference for the flavour develops (Stefurak, Martin, & van der Kooy, 1990). This suggests that the motivational effect conditioned to the first CS becomes paired with the second CS. In the present experiments, some motivational properties may similarly have been acquired by the keylights. These appetitive properties however were largely secondary to the discriminative properties in terms of influencing behaviour. If the value of a conditional reinforcer is not responsible for its behavioural effects, then value may also not be responsible for primary reinforcer effects.

Chapter V

5.1 Experiment 3

Together, Experiments 1 and 2, along with Davison and Baum's (2006; 2010) experiments, demonstrate that choice after a response-contingent non-food stimulus is a function of what that stimulus signals about the time and/or the location of the next food. When the stimulus (uniquely and non-redundantly) signals that the next food is likely to be on the same alternative, preference in a two-alternative procedure will be for the alternative that produced the stimulus. When the stimulus signals that reinforcement is more likely on the other alternative, preference will be towards that not-just-productive alternative.

There was also some suggestion in Experiment 2 that preference after a food was subject to the contingencies of response-contingent event delivery in that post-food period. In Condition 5, when the red-keylight ratio was 9:1, the food ratio was 1:1, and foods and red keylights strictly alternated, choice after a red keylight stabilized at indifference, while choice after a food stabilized at preference for the left. Local preference after a response-contingent event (whether red keylight or food) was a function of the local contingencies in that post-event period.

Further evidence of control by the local contingencies of food reinforcement comes from strict alternation procedures, where a reinforcer from one location is always followed by a reinforcer from a different location. Performance in these procedures is apparently dependent on events or contingencies immediately after a reinforcer delivery. Hearst (1962) for example, reported that over half of his pigeons were more accurate at the intermediate IRIs (2 or 3 s) than at either of the extreme IRIs (1 or 5 s). Williams (1971a; 1971b) found that pigeons could adhere to the strict alternation contingencies when the response requirement was either 15 or 30 responses but not when it was 1 or 5 responses. Both with the long and the short response requirements, the pigeons responded to the just-productive alternative immediately after a reinforcer. However, only when a larger response requirement was in place could they then respond to the other alternative.

Krägeloh, Davison and Elliffe (2005) similarly found that preference to the alternative more likely to provide the next reinforcer was more extreme when this was the just-productive alternative (when the probability of a continuation reinforcer was .5 or greater) than it was the not-just-productive alternative (when the probability of a same-alternative reinforcer was less than .5). Despite this tendency to stay at the just-productive alternative, preference after a reinforcer (as measured by the probability of staying and the preference pulse) was a function of the overall probability of a continuation. Local preference was further towards the not-just-productive alternative when the probability of a continuation was 0 than in any other condition when this probability was greater than 0 (Krägeloh et al.'s Figures 8-10).

Thus, although animals can learn to respond to the not-just-productive alternative when the contingencies favour this, accuracy in these strict alternation procedures appears to be imperfect. An experiment by Davison, Marr and Elliffe (in press) may be informative on why strict alternation is so difficult. In two parts of their experiment, they arranged negative feedback functions such that the behaviour ratio in one IRI determined the relative reinforcer probability in the next IRI: if the behaviour ratio was towards one alternative in an IRI, the next reinforcer was more likely to be on the other alternative. Davison et al. conducted multiple linear regressions between relative choice in an IRI as a function of both relative choice in the prior IRI and the location of the last reinforcer. The effect of the negative contingency was generally small. There was no indication of preference alternation even when the feedback function slope was -3. However there was some effect of manipulating the feedback functions: changing the slope of the feedback function from zero to negative decreased the effect of the previous reinforcer. Additionally, increasing the slope of the negative feedback function (from -1 to -3) further decreased the effect of the prior reinforcer.

Thus, adherence to strict alternation contingencies may be difficult to obtain (Hearst, 1962; Krägeloh et al., 2005; Williams, 1971a, 1971b) because preference in the pre-reinforcer period carries over into the post-reinforcer period. This preference (to the just-reinforced alternative) may interfere with control by local contingencies which require preference to the not-just-reinforced alternative. Hunter and Davison (1985) reported that the reinforcer ratio from 3 to 4 prior sessions still influenced behaviour when the reinforcer ratio changed every session, and Davison and Baum and colleagues have repeatedly found nonzero sensitivity to the prior component's reinforcer ratio at the start of a new component in the frequently-changing procedure (Aparicio & Baum, 2009; Davison & Baum, 2000, 2002; Landon & Davison, 2001). This carryover is obtained despite a clearly discriminable blackout signalling the end of the prior component, and hence that a new reinforcer ratio is in effect. Sensitivity to the previous component's reinforcer ratio is zero from the very first reinforcer delivery only when distinctive discriminative stimuli signal each component reinforcer ratio (Krägeloh & Davison, 2003; Krägeloh et al., 2006). Choice may be generally biased toward repeating the prior response, whether or not that response was just reinforced. Shimp (1966) reported that the probability of making a response on a particular trial was greater if that response had been emitted on the prior trial and Schneider and Davison (2005) reported a bias against response sequences requiring changeovers (e.g., left-right or right-left) when reinforcers were arranged contingent on particular two-response sequences.

One factor that may contribute to this apparent tendency to repeat the prior response may be the changeover contingencies. Choice procedures typically arrange some sort of changeover contingency whereby responses to one alternative cannot be reinforced until some minimal time or number of responses has elapsed since switching from the other alternative. Such changeover contingencies are intended to prevent random responding without regard to the relative reinforcer distribution and thus ensure control by that relative reinforcer distribution (Herrnstein, 1961; Shahan & Lattal, 1998). They are successful in this regard: Increasing the changeover requirement increases measures of sensitivity to reinforcement (Baum, 1982; Davison & Elliffe, 2000; Pliskoff, Cicerone, & Nelson, 1978; Shull & Pliskoff, 1978; Temple, Scown, & Foster, 1995), although there does appear to be a limit to this (Davison, 1991).

An effect of the changeover requirement was reported in Experiment 1: Preference after an unpaired red keylight which signalled nothing about the likely location of the next food was reliably (somewhat) towards the just-productive alternative. This small preference pulse was attributed to the COD: for its duration, responses to the not-just-productive alternative could not be reinforced. A reinforcer could be delivered to the just-productive alternative however (if arranged). Thus, even when the probability of arranging a reinforcer on the just-productive alternative was low, the *local* probability of receiving a reinforcer on that alternative was still higher than the local probability of receiving a reinforcer on the other, not-just-reinforced alternative. Krägeloh and Davison (2003) confirmed the importance of changeover contingencies to the preference pulse: When there were no changeover contingencies and the animals had no prior experience with changeover contingencies, preference pulses were attenuated relative to conditions with active changeover contingencies or a history of changeover contingencies. The local reinforcer ratios reported in Experiment 2 further confirmed that changeover contingencies can change the local obtained reinforcer ratio. Even in conditions where the *arranged* local food ratio in a post-keylight period was 1:1, the local *obtained* food ratio in the first few seconds after a keylight was extreme towards the just-productive alternative. The relative paucity of responses to the not-just-productive alternative meant that a greater portion of the foods received immediately after a response-contingent event were obtained on the just-productive alternative. This local food ratio then drove the local behaviour ratio to more extreme levels. In short, the extreme local behaviour ratio engendered an extreme local reinforcer ratio which entered into a positive-feedback relation with the local behaviour ratio, maintaining extreme local preference.

Experiment 3 investigated the influence of such changeover contingencies on choice in the period immediately after a reinforcer as well as on preference throughout the post-food period. Specifically, Experiment 3 asked what effect the changeover contingencies have on local preference in a strict alternation procedure. The overall reinforcer ratio was the same in all conditions of Experiment 3: 50% of reinforcers were arranged for left-key pecks and the other 50% for right-key pecks. When foods strictly alternated, the probability of a left food was 0 after a left food and was 1.0 after a right food. When foods randomly alternated, the probability of a left food after a left or a right food was .5.

Experiment 3, like Experiment 2 and Davison and Baum's (2006; 2010) experiments, was a two-alternative procedure with a switching key and an explicit changeover response requirement rather than a changeover delay. Thus, only one side key was on at a time and a response requirement needed to be fulfilled in order to switch from responding on one alternative to responding on the other. Use of this response requirement allowed the post-food changeover contingencies to be manipulated by simply varying which alternative came on immediately after food. In some conditions, the just-productive alternative always came on after a food while in others, the not-just productive alternative always came on. In yet another set of conditions, there was no consistent relationship between the alternative that came on after a food and the location of that food: the just-reinforced and not-just-reinforced alternatives were equally likely to come on after a food.

5.2 Method

Subjects

Six homing pigeons, all experimentally naïve at the start of training and numbered 141 through 146 served as subjects. Pigeons were maintained at 85% ±15 g of their of their free-feeding body weights by postsession supplementary feedings of mixed grain when required. Water and grit were freely available in the home cages at all times. The home cages were situated in a room with about 80 other pigeons.

Apparatus

Each pigeon's home cage also served as its experimental chamber. Each cage was 385 mm high, 370 mm wide and 385 mm deep. Three of the walls were constructed of metal sheets and the fourth wall and floor were metal bars. 60 mm above the floor were two wooden perches, one parallel and the other at a right angle to the back wall. Three 20-mm diameter circular translucent response keys were on the right wall. These keys were 85 mm apart and 220 mm above the perches. The keys required a force exceeding approximately 0.1 N to register an effective response when illuminated. The food magazine was also on the right wall 100 mm below the center key. It measured 50 mm high by 50 mm wide and was 40 mm deep. A food hopper, containing wheat, situated behind the magazine, was raised and the magazine was illuminated during food presentations. All experimental events were arranged and recorded on an IBM-PC compatible computer running MED-PC software which was in a room adjacent to the colony room.

Procedure

As these animals were experimentally naïve prior to this experiment, preliminary training, which consisted of magazine-training, autoshaping and VI training, was first implemented. These procedures were identical to those used prior to Experiment 2. The changeover-response requirement was two pecks: The first peck to the centre key turned off the side-key that had been lit, and the second peck to the centre key turned on the other key and turned off the centre key. The first peck to the newly lit side-key turned the centre key back on, once again allowing changeover responses.

Data collection commenced when the pigeons were all responding reliably to the VI 27-s schedule of food reinforcement. In Condition 1, foods randomly alternated across the left and right keys. After any food (from the left or right), and at the start of the session, the next food was randomly allocated to the left or right key with $p = .5$. A probability gate was interrogated every 1 s to determine when food would be scheduled. Once food was set up on an alternative, the next peck to that key turned off that keylight and the red centre key, illuminated the food magazine, and raised the food hopper for 3 s. After food delivery, the hopper was lowered, the magazine darkened and the keylight that had been on prior to the reinforcer (as well as the centre switching key) were illuminated.

Across conditions, whether the foods strictly or randomly alternated was varied as was the alternative that came on after each food. When the foods strictly alternated, a left food was always followed by a right food and vice versa. Which alternative was illuminated immediately after a food was also varied. When the just-reinforced alternative always came on, the events after a food were as described for Condition 1. When the not-just-reinforced alternative always came on, it was the keylight that had not been pecked prior to the last food that came on after the food. In a further set of conditions, the just-reinforced alternative came on after a food with $p = .5$. Table 5.1 presents the sequence of conditions in Experiment 3. Conditions 2, 3, 6 and 11 were not relevant to this experiment and are reported in the next chapter. All sessions lasted 60 minutes or 60 reinforcers, whichever came first and each condition lasted 65 sessions.

Table 5.1

Sequence of conditions, food alternation (strict or random), and the probability that the just-reinforced alternative would be lit after food in Experiment 3.

Cond.	Food alternation	*p*. that the just-reinforced alternative will be lit
1	Random	1.0
4	Strict	1.0
5	Strict	0.5
7	Random	0.5
8	Random	0.0
9 (Replication of Cond. 1)	Random	1.0
10	Strict	0.0
12 (Replication of Cond. 4)	Strict	1.0

5.3 Results

Although behaviour did appear to stabilize faster in Experiment 3 than in Experiment 2, there was still some behaviour change late into some conditions (see Appendix H) again leading to the decision to analyze only the last 20 sessions of each 65-day condition. Using these data, the local response ratio after each response-contingent food was plotted as a function of time since that food (in 2-s bins). These are presented below in Figure 5.1 for conditions which arranged random alternation of foods (Conditions 1, 9, 7 and 8). There are two plots for Condition 7: local preference when the just-productive alternative was illuminated is plotted separately from local preference when the not-just-productive alternative was illuminated. These preference pulses were made in the same way as they were in Experiment 2: left and right responses were tallied in each 2-s time bin following a reinforcer. The log (left/right) response ratio was calculated in each time bin separately following left and right foods and (in Condition 7) following a left keylight onset and a right keylight onset. This was done for each individual subject (Appendix Figures I1-I6). No response ratio was calculated if a time bin contained fewer than 20 responses in total. A value of ±3.5 was used if preference was exclusive to the left or the right in a time bin. The mean log (left/right) response ratio in each time bin was calculated. This is presented below in Figure 5.1.

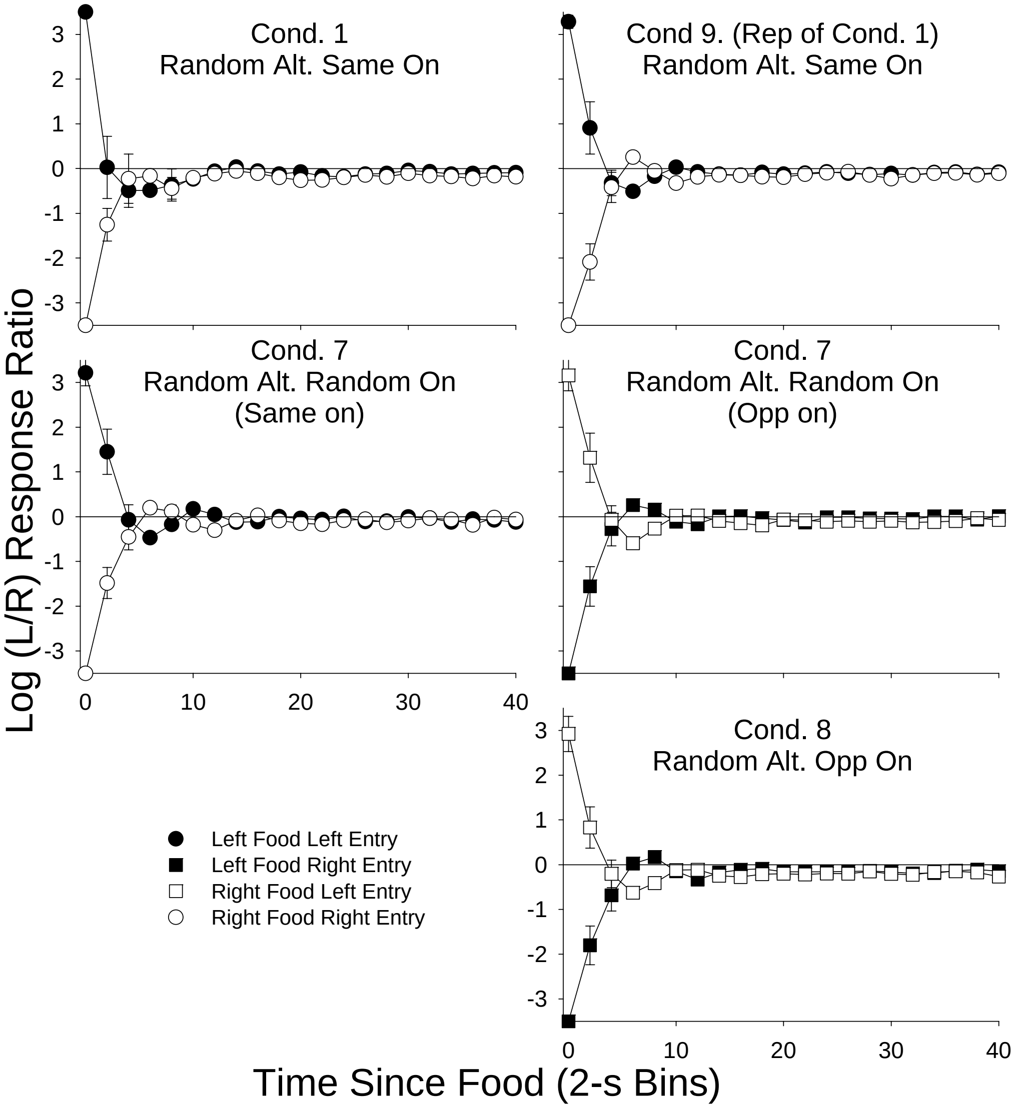


Figure 5.1. Group mean log (L/R) response ratio in each successive 2-s time bin after a food in the conditions of Experiment 3 in which foods randomly alternated across left and right keys. Error bars (1 standard error) are plotted at representative data points, starting at 0 s and increasing in $\log_2$ units. The horizontal line is at 0.

Local preference after a food in Condition 1 was very similar to local preference after a food in Condition 9 suggesting no effect of extensive experience on local preference after a reinforcer. Generally, the group preference pulses were representative of the individual subject preference pulses (Appendix Figures I1-I6). In all of these random-alternation conditions, preference after a reinforcer was towards the alternative illuminated immediately after that reinforcer. This was the case whether the just-productive alternative was illuminated (Conditions 1 and 9 always and about half of the time in Condition 7), or whether the not-just-productive alternative was illuminated (all of Condition 8 and the other half of Condition 7). There was no apparent effect of the location of the last food. Figure 5.2 presents local preference after a food in conditions where foods strictly alternated. These were calculated in the same way as the preference pulses in Figure 5.1. The individual-subject preference pulses are presented in Appendix Figures I7-I13.

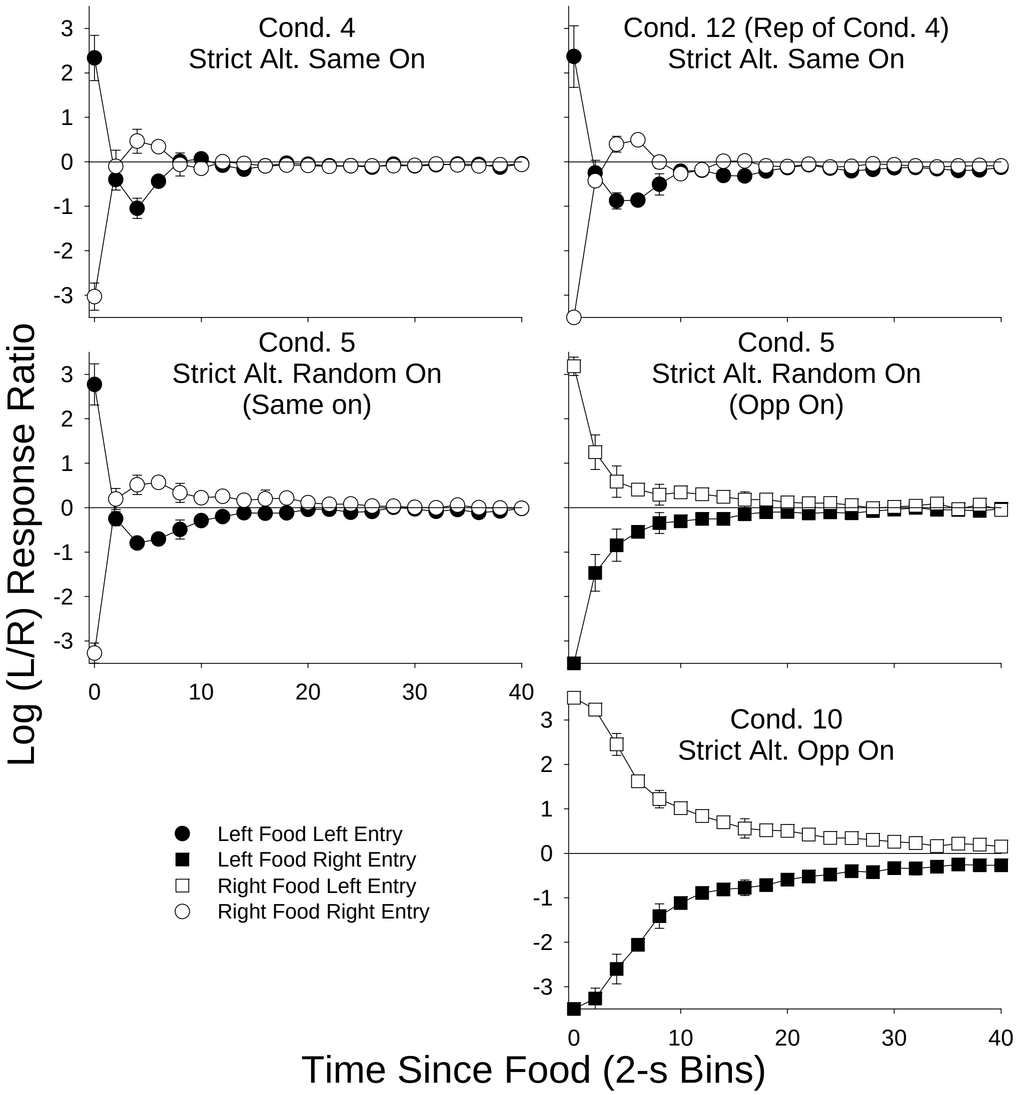


Figure 5.2. Group mean log (L/R) response ratio in each successive 2-s time bin after a food in the conditions of Experiment 3 where foods strictly alternated between left and right keys. Error bars (1 standard error) are plotted at representative data points, starting at 0 s and increasing in $\log_2$ units. The horizontal line is at 0.

Once again, the replication indicated that preference pulses did not change throughout the experiment. Local preference after a food in Condition 12 was nearly identical to local preference after a food in Condition 4. The most notable feature of Figure 5.2 is that there was always some detectable preference for the not-just-productive alternative in all of these strict alternation conditions. The extent of this preference was determined by the keylight that came on after a food. Preference was more extreme in Condition 10 (when the not-just-productive alternative always came on after a reinforcer) than in any other condition. In Condition 5 (when the just-reinforced alternative came on half of the time and the not-just-reinforced alternative came on the other half), preference was more extreme when the not-just-reinforced alternative came on (middle right panel). The effect of illuminating the not-just-productive alternative immediately after a reinforcer was largely constrained to the period immediately after a reinforcer — after this initial period, local preference in Condition 5 was similar whether the just-reinforced, or the not-just-reinforced alternative was illuminated.

For all individual subjects, preference for the not-just-productive alternative was more extreme when that alternative always was illuminated immediately after a reinforcer (Condition 10) than when it was only illuminated 50% of the time (Condition 5). Additionally, group preference for the not-just-productive alternative was more extreme when the just-productive-alternative was illuminated in Condition 5 than in Conditions 4 and 12 (when the just-productive alternative was always illuminated). Although this was not clearly demonstrated in all individual subjects, it was clearly present in 3 of the 6 individuals (with a 4th perhaps also trending in this direction). Thus local post-food preference was affected by at least 3 factors in this experiment: first, the local contingencies of reinforcement (strict or random alternation); second, the alternative that was illuminated after the most recent food and; third, the overall probability that the not-just-productive alternative would be illuminated after a food. When this probability was high (1.0 in Condition 10), preference to the not-just-productive alternative was extreme; when it was low (0.0 in Conditions 4 and 12) preference was close to indifference, and when this probability was intermediate (.5 in Condition 5) preference was also intermediate. It may be that the local contingencies of reinforcement were better learned when there was a high probability that the post-food illuminated alternative would provide the next food. This was evaluated more objectively by calculating cumulative log response ratios for the entire period following each reinforcer-keylight onset combination in each of the strict alternation conditions. The group log response ratios are presented below in Table 5.2.

Table 5.2

Group mean log(L/R) response ratio for the entire period following a food in all Experiment 3 conditions with strict alternation of reinforcers. One standard error is presented in parentheses.

Cond.	Left Food, Left Onset	Left Food, Right Onset	Right Food, Left Onset	Right Food, Right Onset
4	-0.15 (0.07)			-0.01 (0.10)
12	-0.26 (0.08)			-0.05 (0.05)
5	-0.20 (0.09)	-0.27 (0.11)	0.24 (0.11)	0.16 (0.12)
10		-0.65 (0.11)	0.60 (0.11)	

In order to evaluate the effects of foods in each condition, three 2x2 factorial ANOVAs were conducted with condition and food-location (left vs. right) as the factors. Two of these ANOVAs compared the log behaviour ratio after left and right foods when the just-productive alternative came on (Condition 4 vs. Condition 5 and Condition 12 vs. Condition 5). The third compared these log behaviour ratios when the not-just-productive alternative came on after a food (the other half of Condition 5 vs. Condition 10). In all of these ANOVAs ($\alpha < .05$), there was a significant main effect of food's location (left vs. right) and no significant main effect of condition. There was no significant interaction when the just-productive alternative was illuminated (half of Condition 5 vs. either Condition 4 or Condition 12). There was however a significant food-condition interaction when the not-just-productive alternative was illuminated (Condition 10 vs. the other half of Condition 5, $\alpha < .01$). These results indicate that in the strict alternation conditions, the log behaviour ratio after left and right foods were always reliably different from one another. This difference was enhanced when the just-productive alternative always came on after a food relative to when it only came on after a food 50% of the time. There was no enhancement however when the just-productive alternative only sometimes came on after a food relative to when it always came on after the food.

The greater preference (in the strict alternation conditions) for the not-just-productive alternative when that alternative was illuminated may have been an artefact due to the changeover requirement. The 2 responses to the centre key required to changeover from one key to the other took time to emit and delayed subsequent responses. Any local increase in preference to the not-just-productive alternative would thus be shifted along the x-axis by the time taken to complete the changeover requirement. Figures 5.3 and 5.4 plot the preference pulses as in Figures 5.1 and 5.2 respectively, but with all times taken to make any changeover responses removed (the individual subject preference pulses are presented in Appendix Figures I13-I24).

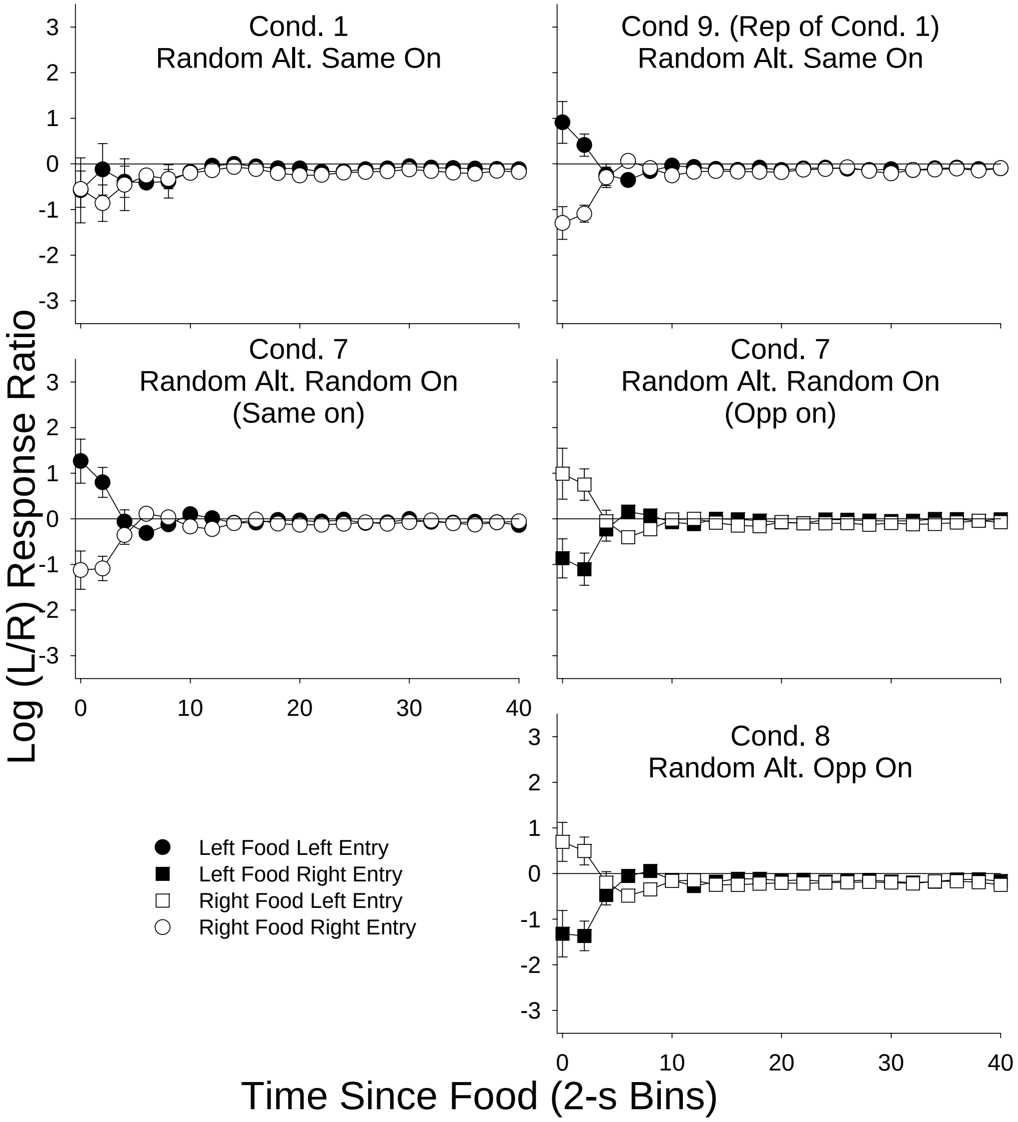


Figure 5.3. Group mean log (L/R) response ratio in each successive 2-s time bin after a food in the conditions of Experiment 3 where foods randomly alternated across left and right. All changeover time was removed from these plots. Error bars (1 standard error) are plotted at representative data points, starting at 0 s and increasing in $\log_2$ units. The horizontal line is at 0.

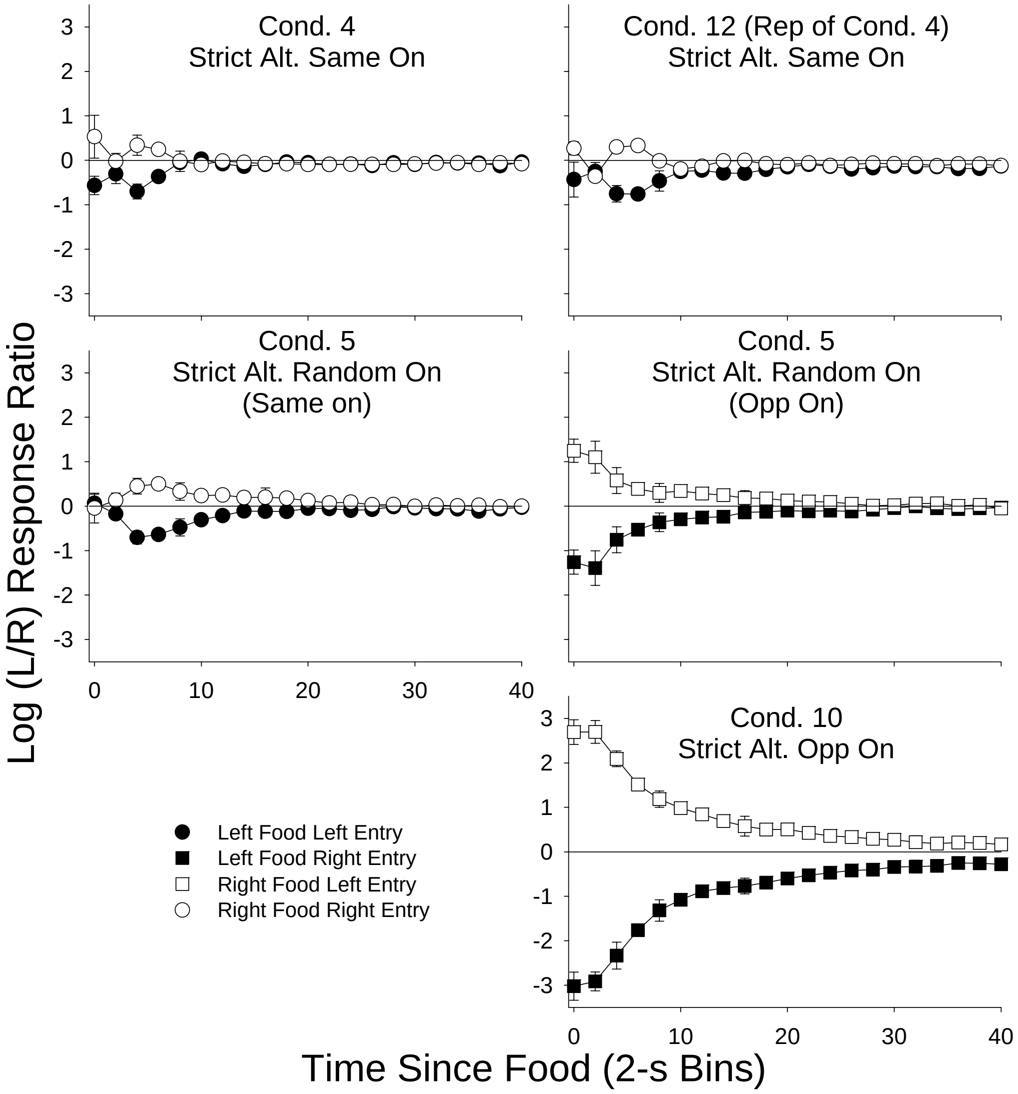


Figure 5.4. Group mean log (L/R) response ratio in each successive 2-s time bin after a food in the conditions of Experiment 3 where foods strictly alternated across left and right. All changeover time was removed from these plots. Error bars (1 standard error) are plotted at representative data points, starting at 0 s and increasing in $\log_2$ units. The horizontal line is at 0.

In both the random-alternation conditions (Figure 5.3) and the strict-alternation conditions (Figure 5.4), removing all changeover times made initial preference less extreme (relative to Figures 5.1 and 5.2) but had no other effects. About 6 s from the time of the food, there were no differences between the data shown in Figures 5.1 and 5.3 and between Figures 5.2 and 5.4, indicating that removing the time taken to complete changeovers only affected the start of the preference pulse, and that this effect was restricted to simply making the preference pulse less extreme. Because this was the case in both the random- and strict-alternation conditions, this result indicates that the time taken to complete a changeover was not responsible for the greater preference for the not-just-reinforced alternative when this alternative was illuminated after a food. Instead, there was apparently some tendency to stay at the first-illuminated alternative in every condition, regardless of the local probability of a reinforcer on this alternative.

The post-food changeover contingencies and probability of a same-alternative reinforcer may also influence the probability of staying at the just-productive alternative. Figure 5.5 presents, for each individual subject in all of the random alternation conditions, the probability of staying at the just-productive alternative after a food reinforcer.

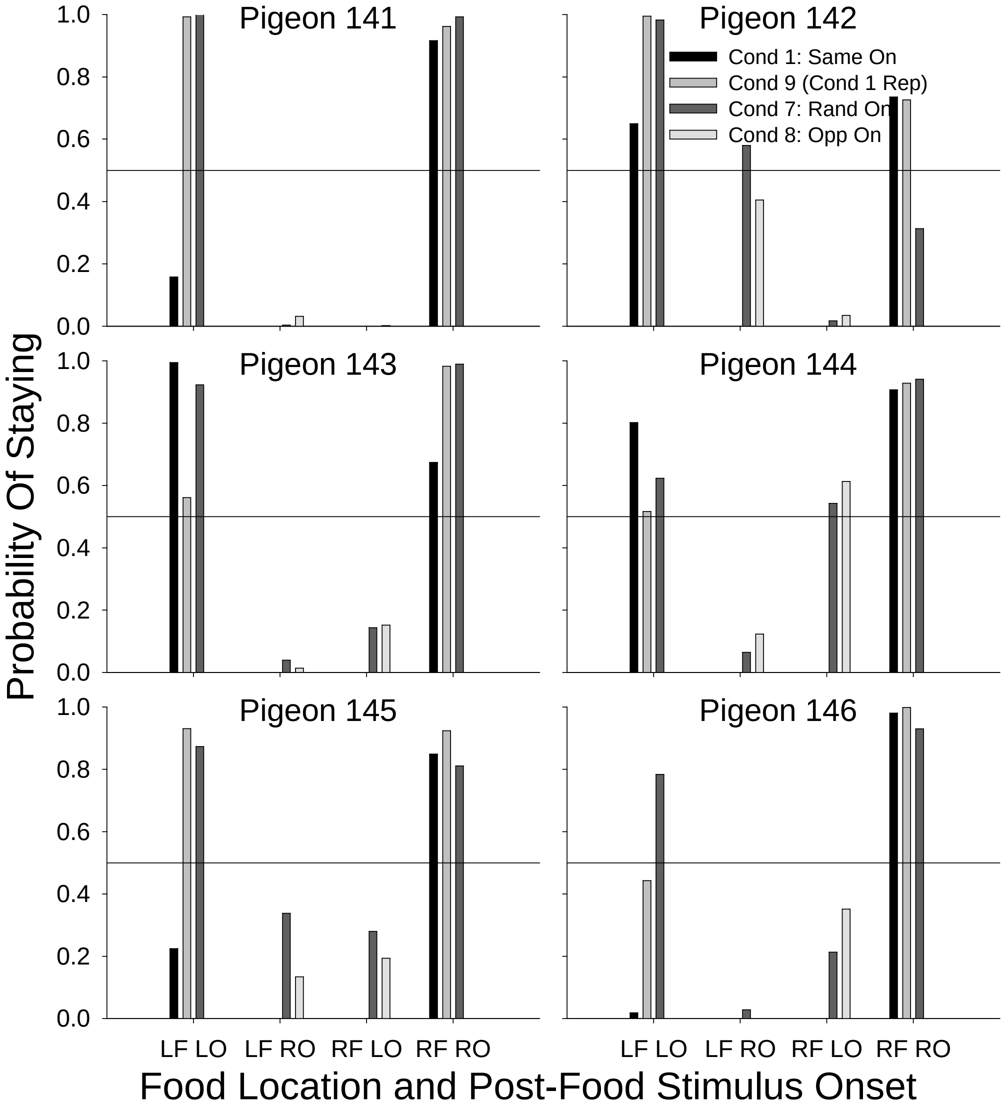


Figure 5.5. The probability of staying at the just-reinforced alternative after left and right foods followed by a left or right stimulus-onset, for each individual subject in each condition with random alternation of foods. On the x-axis, LF denotes left food, RF right food, LO left onset and RO right onset. Each combination thus denotes a unique food-onset combination. There were no instances of LF LO or RF RO in Condition 8, and no instances of LF RO or RF LO in Conditions 1 and 9.

Figure 5.5 demonstrates that, in the random-alternation conditions, the probability of staying at the just-reinforced alternative was rather variable and was apparently heavily influenced by position biases. In Condition 1, Pigeons 141, 145 and 146 were likely to stay after a right reinforcer and likely to switch after a left reinforcer. This was partially true for Pigeon 143, while Pigeons 142 and 144 were about equally likely to stay after left and right foods. In the replication of Condition 1 (Condition 9), these position biases were somewhat ameliorated. All 6 of the pigeons were now likely to stay after a right reinforcer, and 5 of the 6 animals were likely to stay after a left reinforcer. In Condition 7, Pigeons 141, 143, 145 and 146 were very likely to stay at the just-reinforced alternative when it was illuminated immediately after the food, and were very likely to switch when the not-just-reinforced alternative was illuminated after a food (the probability of staying was very low). Pigeon 142 was more likely to stay after a left food regardless of whether the left or right key came on. Pigeon 144 was likely to stay in all cases except when a left food was followed by a right onset. In Condition 8, the probability of staying was generally low for all subjects, indicating that they were all likely to respond to the first-illuminated alternative rather than switching to the alternative that produced the last food. Only in 1 of 12 cases was the probability of staying at the just-productive alternative higher than .5 in Condition 8.

Thus, although there did appear to be rather a lot of between-subject variability, as well as apparent position biases, the probability of staying at an alternative was generally a function of the alternative that came on after the food: if the just-productive alternative came on after the food (LF LO and RF RO), the probability of staying was generally higher than .5. If the not-just-productive alternative came on (LF RO and RF LO), then the probability of staying was generally less than .5. Any exceptions to this were likely due to position biases and were never symmetrical across the two alternatives (i.e., no animal was ever more likely to switch immediately after a left food, left onset and after a right food, right onset). Thus, when the foods randomly alternated, the location of the first visit (just-productive alternative versus not-just-productive alternative) was likely to be to the alternative that was illuminated immediately after a reinforcer. This conclusion is consistent with the conclusion that followed the preference pulse analyses. Figure 5.6 presents the same plots for conditions arranging strict alternation of foods.

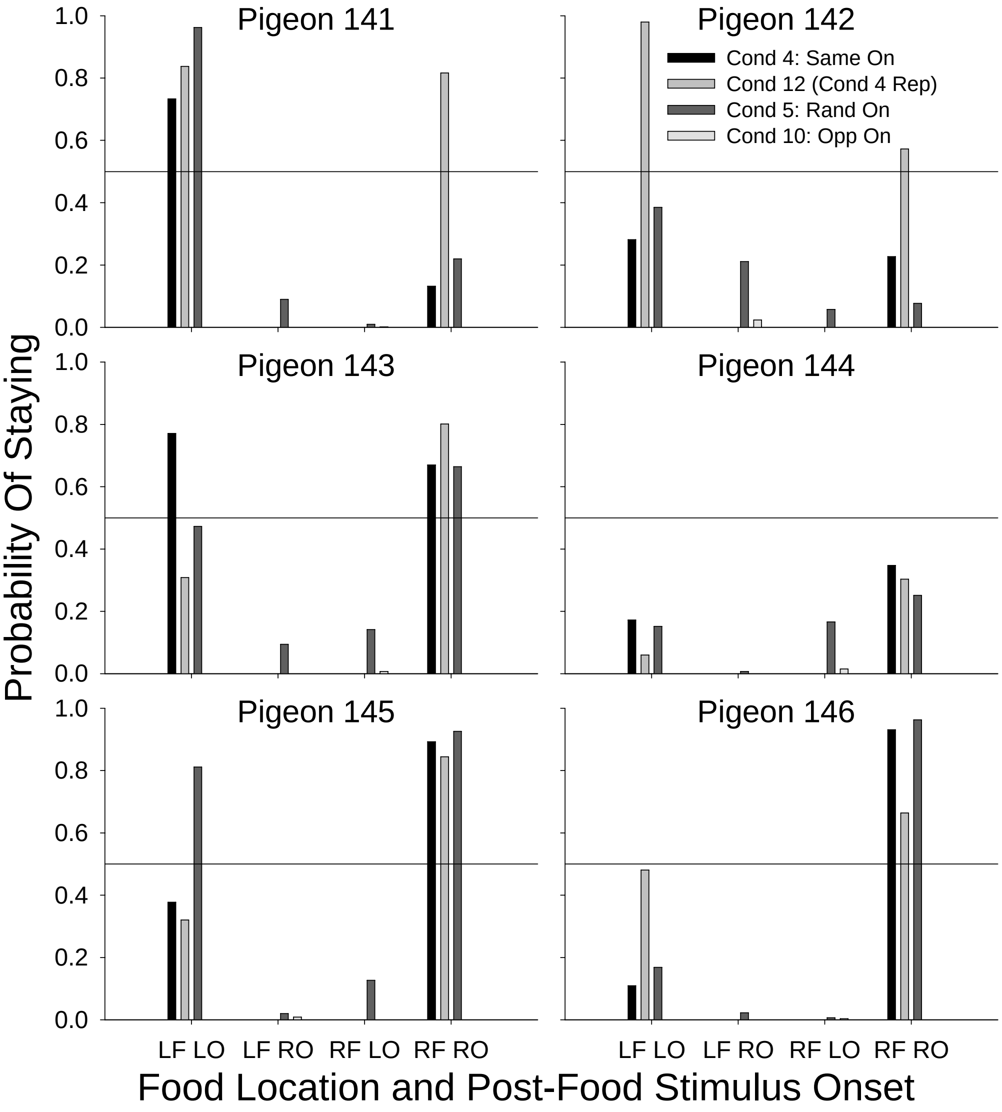


Figure 5.6. The probability of staying at the just-reinforced alternative after left and right foods followed by a left or right stimulus-onset, for each individual subject in each condition with strict alternation of foods. On the x-axis, LF denotes left food, RF right food, LO left onset and RO right onset. Each combination thus denotes a unique food-onset combination. There were no instances of LF LO or RF RO in Condition 10, and no instances of LF RO or RF LO in Conditions 4 and 12.

Although the probability of a reinforcer at the just-productive alternative was always 0 in the conditions shown in Figure 5.6, the probability of staying at that alternative was often greater than .5. In Condition 4 the probability of staying was greater than .5 In 5 of 12 cases. In Condition 12 the probability of staying was greater than .5 In 7 of 12 cases. In these conditions, despite responses to the just-productive alternative never being reinforced, the first visit was often to that alternative. In Condition 5, when the post-food stimulus onset was random, the probability of staying remained high when that alternative came on immediately after the food in 5 of 12 cases. When the not-just-productive alternative came on, the probability of first responding to the just-productive alternative was always very low, whether that alternative always came on (Condition 10) or came on about half of the time (Condition 5). Thus, Figure 5.6 demonstrates that, although the preference pulses suggested control by the local contingencies of reinforcement in these strict alternation conditions, there was no strong control over the probability of staying. This probability of staying was still largely controlled by the alternative that came on immediately after the food.

Reinforcer effects are not restricted to the period immediately after their delivery and before the delivery of the next reinforcer. Reinforcers also have effects extending across multiple reinforcer deliveries. The best known of these is generalized matching in which the long term allocation of behaviour across two alternatives is a function of the long term aggregation of reinforcers across those alternatives. At a somewhat less extended level, Landon et al. (2002) showed, in a steady-state concurrent schedule procedure, that reinforcers going as far back as 8 had a measurable effect on behaviour. The analysis used by Landon et al. was carried out on the present data. This analysis investigated preference as a function of the sequence of the preceding 8 foods and thus could only be done on conditions arranging random alternation of foods. First, the log (left/right) response ratio was calculated for each of the 256 possible 8-food sequences of left and right foods. Then, the contribution of each of the preceding 8 reinforcers to the response ratio was calculated using the general linear model:

$$\log\left(\frac{B_L}{B_R}\right) = \log k + \sum_{j=0}^{7} \begin{Bmatrix} R_j = L : +b_j \\ R_j = R : -b_j \end{Bmatrix} \qquad \text{(Equation 5.1)}$$

where the subscript j indicates the reinforcer lag so that R_0 is the most recent reinforcer and R_7 is the 8th reinforcer back. The effect of a reinforcer at a given lag is measured by the coefficient b_j. If a reinforcer was from the left, this b_j was multiplied by +1; if the reinforcer was from the right, it was multiplied by -1. The constant log k is the residual preference not due to any of the previous 8 reinforcers. These values were calculated by finding the best-fitting least-squares estimates of b_j and log k using Excel's solver function. These b_j and log k values are presented in Figure 5.7 for each individual subject in each of the random-alternation conditions.

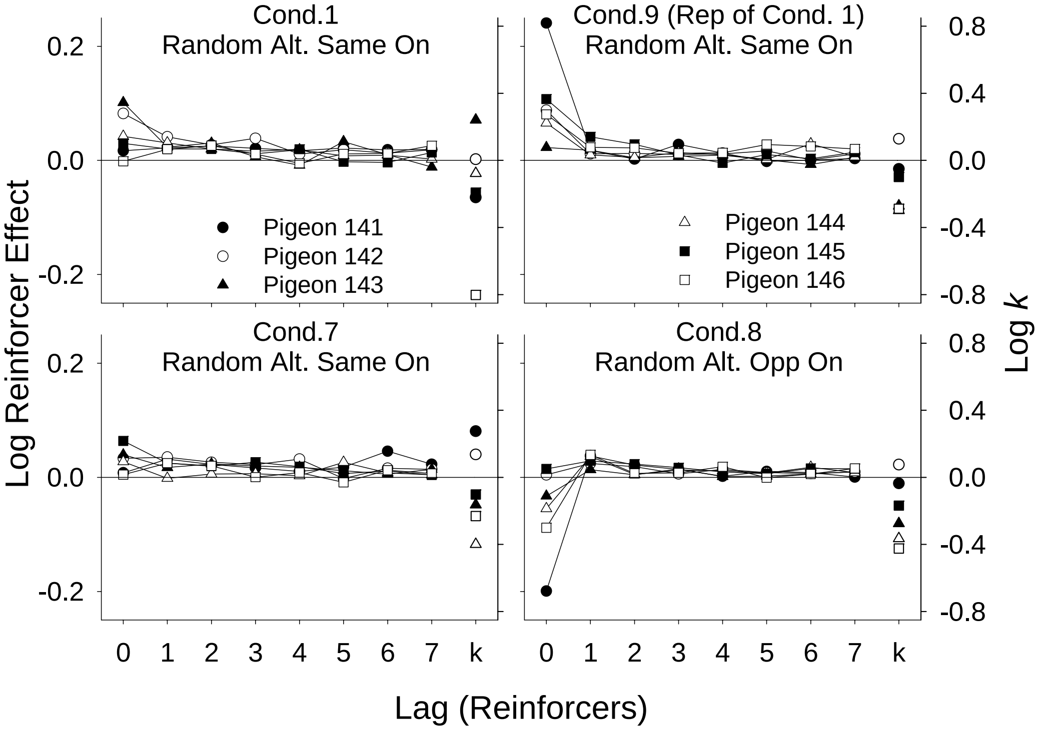


Figure 5.7. Lag reinforcer effects and log *k* (preference not due to any of the 8 preceding reinforcers) for all individual subjects in all random-alternation conditions in Experiment 3. The lag reinforcer effects (connected by a straight line) are relative to the left y-axis and the log *k* (single point, not connected to the others) is relative to the right y-axis.

The effect of any reinforcer, even the most recent one, was always rather small. Landon et al. (2002) also reported small reinforcer lag effects. The reinforcer lag effects reported here appear consistent with those obtained in Landon et al.'s Condition 1, which also arranged a 1:1 food ratio. In the present experiment, the reinforcer lag effects in Condition 1 were only imperfectly replicated in Condition 9 (unlike the preference pulse results which were much more closely replicated). The most recent reinforcer had a larger effect in Condition 9 than in Condition 1 (though not for all subjects) and log *k* (inherent bias or bias due to reinforcers further back than the 8th) generally (again though not for all subjects) decreased. The effect of the most recent reinforcer was rather smaller in Condition 7 when either alternative could be illuminated after a reinforcer. Also, the log reinforcer effect for the first reinforcer back was negative in Condition 8 when the not-just-reinforced alternative always came on after a reinforcer. This indicates that preference was shifted away from the alternative that provided the most recent reinforcer in that condition (consistent with the preference pulses). This apparent preference for the not-just-productive alternative may be better understood as a preference to the first-lit alternative. The preference to the just-productive-alternative in Condition 9 (and to a lesser extent in Condition 1) when the just-productive alternative was illuminated after a reinforcer may also be better understood as a preference to the first-illuminated alternative.

Although the reinforcer lag effect did drop off somewhat after the first reinforcer, it consistently remained greater than 0, indicating that reinforcers beyond the most recent one had a small effect on current preference. Additionally, while the log reinforcer effect for the most recent reinforcer in Condition 8 was negative, the log reinforcer effect for older reinforcers was always greater than 0 in all conditions, including Condition 8. Reinforcers prior to the most recent one appeared to shift preference towards the alternative that provided that reinforcer in all of the conditions of Figure 5.7, regardless of what alternative was illuminated after a food. Figure 5.8 presents a similar lag analysis except that preference is a function of the sequence of preceding keylight onsets rather than reinforcer locations. This analysis could only be done on conditions arranging random alternation of the post-food keylight onset (Condition 5 which arranged strict alternation of foods and Condition 7 which arranged random alternation of foods).

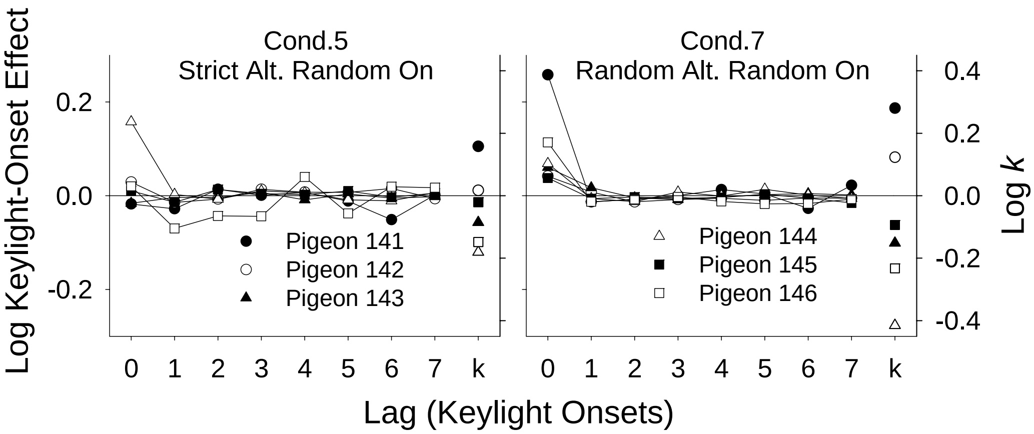


Figure 5.8. Lag keylight-onset effects and log *k* (preference not due to any of the 8 preceding keylight onsets) for all individual subjects in Conditions 5 and 7. The lag keylight-onset effects (connected by a straight line) are relative to the left y-axis and the log *k* (single point, not connected to the others) is relative to the right y-axis.

Figure 5.8 shows that the location of the last keylight onset had a rather large effect, but that the location of keylight onsets prior to the most recent had no effect: although the lag effect of the most recent keylight onset location was always reliably greater than 0, the lag effect of keylight onsets prior to the most recent was always very close to 0. The effect of earlier keylight onsets was lower than the effects of previous foods (Figure 5.7) which although small, were consistently greater than 0. Thus, while the most recently illuminated alternative had a rather large influence on current behaviour (demonstrated in the preference pulses), the alternatives lit immediately after earlier reinforcers had minimal effect. Conversely, while there was no apparent local effect of the most recent reinforcer location, there was some indication of small, long-lasting and cumulative effects of prior reinforcers.

Reinforcer-sequence effects can also be investigated by plotting reinforcer trees. In these plots, the log response ratio in successive IRIs is plotted as a function of a selected sequence of previous reinforcers. As these plots can quickly get out of hand, Figure 5.9 only plots sequences of continuations and single discontinuations for the group mean. The individual subject plots are presented in Appendix Figures I25-

I30. Once again, these plots could only sensibly be created for conditions where the foods randomly alternated.

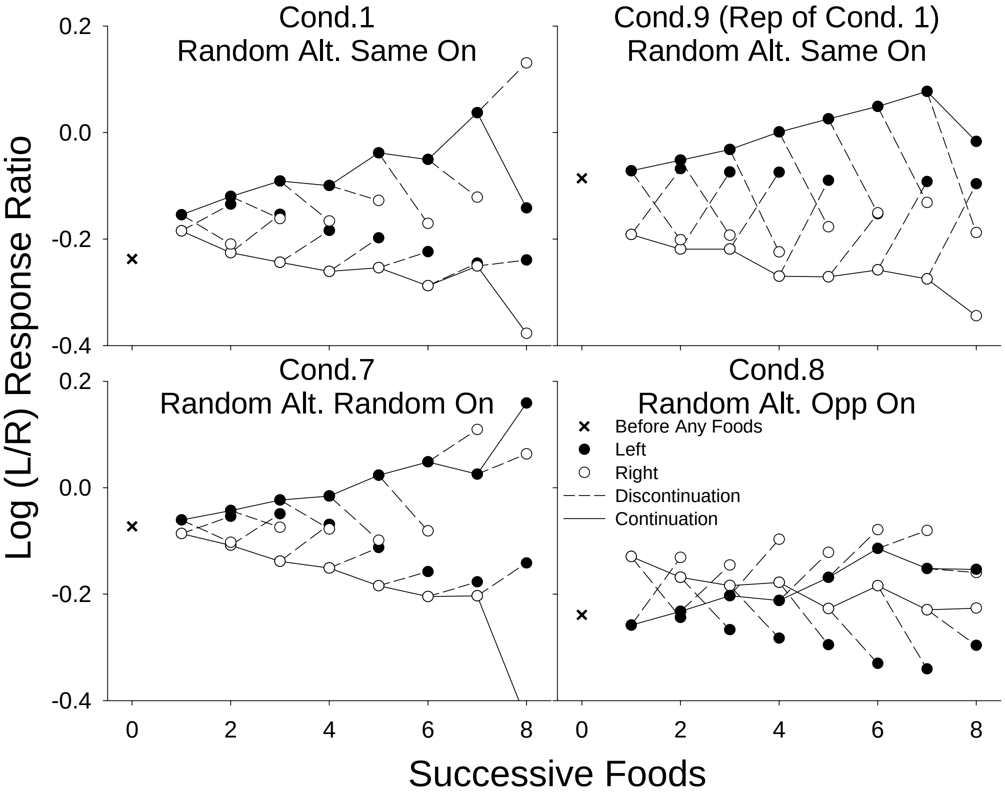


Figure 5.9. Group mean log (L/R) response ratio in successive inter-reinforcer intervals as a function of the location of successive reinforcers in the random alternation conditions.

Figure 5.9 replicates a number of findings previously reported in both frequently changing and steady-state environments. Conditions 1 and 9 were most similar to more typical concurrent schedule procedures: in those conditions, the foods randomly alternated (overall food ratio = 1:1), and the just-productive alternative was always illuminated immediately after a reinforcer. Successive same-alternative foods (continuations) progressively shifted preference to the alternative providing those foods. Discontinuations (reinforcers from the alternative that did not provide the prior reinforcers) shifted preference further than did continuations. Having said this, discontinuations did not completely reverse the extreme preference engendered by prior continuations: There was a detectable residual effect of reinforcers prior to the most recent one and preference after a discontinuation was clearly influenced by the number of preceding other-alternative continuations. These general trends and patterns were obtained both in Condition 1 and its replication Condition 9, although reinforcer effects were generally larger in Condition 9. The preference tree for Condition 7 was also largely consistent with these conclusions. Thus, even when the alternative illuminated after a reinforcer was completely random, the preference-increasing effects of reinforcers cumulated across successive reinforcer deliveries.

The preference tree from Condition 8, when reinforcers randomly alternated and the not-just-productive alternative was always illuminated after a food, was rather unlike any other preference tree in that there was a distinct cross-over: preference after four or fewer foods from a single source was towards the not-just-productive alternative (the alternative that was illuminated immediately after those reinforcers). However, there was also a general trend for preference to move towards the alternative providing the (continuation) reinforcers. After five or more same-alternative reinforcers preference was towards the alternative providing those reinforcers. As in the other conditions of Figure 5.9, even though reinforcers were arranged without regard to any of the prior reinforcers, behaviour was still a function of the location of those reinforcers.

Preference prior to a reinforcer carries over into the period beyond that reinforcer regardless of the post-food changeover contingencies. The effects of continued same-alternative reinforcers, while small, were cumulative and thus visible after a series of reinforcers from the same alternative. There is still some question as to whether the increased preference for the alternative producing the reinforcers was due to simply the strengthening effects of the previous reinforcers, or whether the animals were, in some sense, predicting that the alternative providing these reinforcers is locally rich. Thus, the question becomes: Is the pigeon predicting that the future will be the same as the past (reinforcers will continue to come from the same alternative), or is preference being driven towards the alternative providing successive continuation reinforcers? Figure 5.10 examined this by plotting preference as a function of successive discontinuations. Sequences of strictly-alternating reinforcers were selected from the random-alternation conditions and are presented as the trees above.

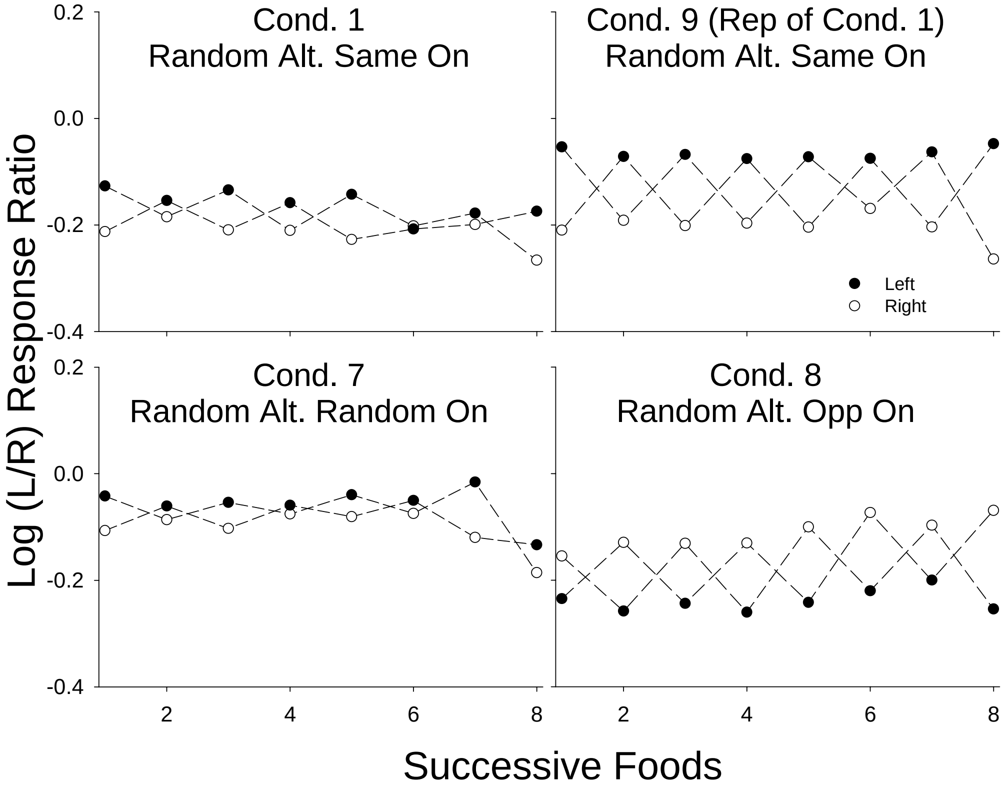


Figure 5.10. Group mean log (left/right) response ratio in the period between successive discontinuations in the conditions of Experiment 3 which arranged random alternation of foods.

Preference was largely unchanged by the number of preceding alternation reinforcers in the group mean data above or in the data of any of the individual subjects in Appendix Figures I31-I36. There was no tendency for choice to be further towards the left after a right food or further towards the right after a left food as the number of preceding alternations increased in any condition. A discontinuation shifted preference away from the extreme levels engendered by previous continuation(s) regardless of the preceding sequence length. Thus, the preference trees in Figure 5.9 were not likely due to the pigeons predicting that the immediate future would be like the immediate past (that the same alternative would continue to provide reinforcers). Instead, there appears to be some tendency for preference to shift towards the alternative that provided the most recent reinforcers.

Preference trees can also be plotted as a function of successive post-food keylight onsets. These are presented in Figure 5.11 for the group data for the two conditions which arranged random alternation of the post-food keylight onset. The individual-subject graphs are presented in Appendix Figures I37-I42.

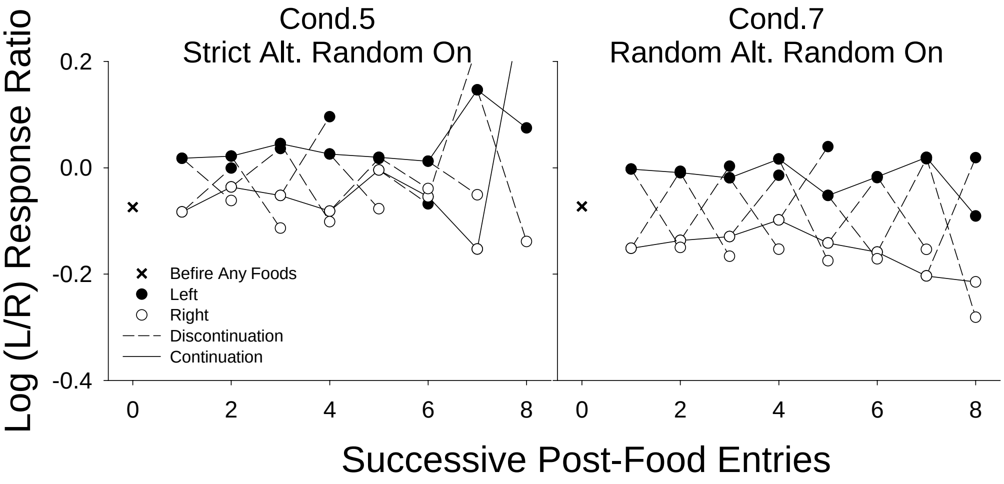


Figure 5.11. Group mean log (L/R) response ratio after a series of reinforcers followed by the same keylight onset or by a keylight onset which differed from the preceding keylight-onset sequence for all of the conditions of Experiment 3 where the post-food keylight onset was random.

Generally, preference was towards the alternative which was illuminated after the most recent reinforcer in both Condition 5 and Condition 7. This is consistent with the preference pulses. The functions were largely flat in both conditions, and there were no trends for preference to increase as an alternative was illuminated after an increasing number of reinforcers. This is consistent with the lag analysis. Successive same-alternative stimulus illuminations, unlike successive same-alternative foods, had no cumulative effects. Also, preference after a stimulus onset from the alternative which did not provide the preceding sequence of stimulus onsets (a stimulus discontinuation), was apparently equal to preference after a stimulus onset from that alternative preceded by same-alternative onsets. The complete lack of control by successive same-alternative post-food stimulus onsets is perhaps better demonstrated in the individual subject plots (Appendix Figures I37-I42). These plots are even less orderly than Figure 5.11, and demonstrate the lack of control by successive post-food stimulus onsets. Thus, while food's location had cumulative and long-lasting effects on behaviour, the effect of the stimulus-onset was constrained to the local, immediate level.

5.4 Discussion

Reinforcers in Experiment 3 were always equally distributed across the two alternatives at the long-term level. In half of the conditions, the location of each food was unknowable prior to its being delivered: a left food could be followed by a left or a right food with equal probability. In the other conditions, the location of each food could be perfectly predicted: a left food was always followed by a right food and a right food was always followed by a left food.

Previous studies using strict alternation procedures have found less-than-perfect accuracy despite the fact that the location of each food could be perfectly predicted knowing only the location of the last food. Krägeloh et al. (2005) reported some asymmetry between conditions arranging a low probability of a continuation and conditions arranging a high probability of a continuation. Preference pulses were further towards the just-productive alternative when the probability of a continuation was 1 than they were towards the not-just-reinforced alternative when this probability was 0. Additionally, when the probability of a continuation reinforcer was .5, preference was reliably towards the just-productive alternative, and the probability of staying at the just-productive alternative was greater than .5. Landon, Davison and Elliffe (2003b) also found local preference to the just-productive alternative in a steady-state procedure with a 1:1 left: right reinforcer ratio. Although Davison et al. (in press) did not arrange a strict alternation procedure as such, the negative relation between the response ratio in one IRI and the reinforcer ratio in the next IRI meant that reinforcers in two phases of their experiment were more likely to occur on the not-just-productive alternative. They also reported difficulty in getting their subjects to conform to the alternation contingencies.

Thus, learning to respond to the not-just-reinforced alternative may be rather difficult (though not impossible). One reason for this difficulty may be the changeover contingencies effective at all times, including the period following a reinforcer. These changeover contingencies were identified as the most likely reason for the preference pulses after the unpaired stimuli in Experiment 1: no reinforcers could be obtained from the not-just-productive alternative for the duration of the changeover period, and this may have biased choice in that period to the just-reinforced alternative. Although perhaps unlikely, there is still some nonzero probability of a reinforcer being collected on the just-productive alternative immediately after a reinforcer. Experiment 2, and Krägeloh and Davison's (2003) results, confirmed the importance of changeover contingencies on local preference after a response-contingent event.

Experiment 3 examined this potential source of bias by varying the alternative from which the animal must change over. A changeover-ratio procedure was used to explicitly do this: In some conditions the just-productive-alternative always came on; in other conditions the not-just-reinforced alternative always came on, and in yet other conditions the alternative to come on immediately after a reinforcer was unrelated to the location of that reinforcer. When reinforcers randomly alternated, preference immediately after a single reinforcer was towards the alternative that came on after that reinforcer (not towards the alternative that delivered that reinforcer). This suggests that the preference pulse, which may appear to be a demonstration of a typical reinforcer effect, may be better understood as a tendency to respond to the alternative momentarily more likely to provide the next reinforcer. The foods in

Experiment 3 signalled that food was forthcoming, just as the red and green keylights did in Experiment 2. When foods randomly alternated between left and right keys, the next food was as likely to be on the just-productive alternative as on the not-just-productive alternative. However because of the changeover contingencies, an arranged food could only be obtained immediately on the alternative that was illuminated immediately after that keylight. Although a reinforcer could be arranged on the other alternative, it could not be obtained until the changeover requirement was completed.

Although local preference was towards the alternative that was illuminated immediately after a reinforcer, the effects of the post-food changeover contingencies were apparently short-lived. When the post-food illuminated alternative was randomly varied, there was no suggestion of an increase in preference to an alternative with increasing numbers of post-food illuminations, under either strict or random reinforcer alternation (Figures 5.8 and 5.11). This is in contrast to the more standard reinforcer lag effects and preference trees (Figures 5.7 and 5.9) which showed typical increases in local preference across successive same-alternative reinforcer deliveries, as well as residual effects of reinforcer sequences after a single discontinuation. Perhaps most compelling of all was the preference tree in Condition 8 which arranged random alternation of reinforcers, and that the not-just-reinforced alternative always came on after a reinforcer. Although the preference pulse was towards the not-just-reinforced alternative in this condition (the alternative that was illuminated after the reinforcer), preference continually increased towards the alternative providing the continuation reinforcers, not to the alternative that came on after those reinforcers. The effect of reinforcer location was not clearly visible after a single reinforcer but was cumulative and long-lasting and became evident after a number of same-alternative reinforcers. Thus, there is a preference-increasing effect of reinforcers that is not attributable to the changeover contingencies.

What of the strict alternation conditions? Strict alternation has proven difficult to obtain, though not impossible. Generally, behaviour tends to be roughly in accordance with strict-alternation contingencies, although there is some tendency to repeat the just-reinforced response. In a strict alternation condition, Krägeloh et al. (2005) found that although local preference was initially towards the just-reinforced alternative, the general thrust of preference was towards the not-just-reinforced alternative. Hearst (1962) found that a U-shaped relation described accuracy as a function of IRI duration in a discrete-trials alternation procedure. Williams (1971a; 1971b) also reported a failure to perform above chance in a discrete trials alternation procedure when less than 15 responses were required, although the task could be completed easily when 15 or more responses were required. This result was attributed to some tendency to respond to the just-reinforced alternative immediately after a reinforcer and to only later be able to respond to the other alternative. Such a tendency to stay at the just-reinforced alternative was also found here in the strict alternation conditions when the just-reinforced alternative came on after the reinforcer (Conditions 1 and 9 and some reinforcers in Condition 5; Figure 5.6). Preference then moved towards the not-just-reinforced alternative before settling at the long-term level of indifference (Figure 5.2). Such a change in preference was not seen when the not-just-reinforced alternative came on after the reinforcer (Condition 10, and the other part of Condition 5). After those reinforcers, initial preference was always towards the not-just-reinforced alternative.

The initial preference to the just-reinforced alternative previously reported may also be better attributed to the changeover contingencies in those experiments, rather than to any drive to repeat the just-reinforced response. This readily explains the local post-food preference in Landon et al.'s (2003b) and Krägeloh et al.'s (2005) experiments which were free-operant two alternative concurrent schedule procedures with typical changeover requirements. Davison et al. (in press) did not arrange any changeover requirement, and no such requirement existed in the discrete trials procedures described above (Hearst 1962; Williams 1971a, 1971b). In these procedures, the animal's position within the chamber and in reference to the response manipulada may mimic the role of the changeover contingencies (Aparicio & Baum, 1997; Baum, 1982). An initial tendency to stay may be a function of bodily position: immediately after a reinforcer the animal is likely to be physically closer to the last manipulandum responded to and thus, may respond to that alternative at shorter times after the reinforcer. Hearst reported that his pigeons appeared to station themselves in front of the correct alternative throughout the IRI. When this IRI was shortest, the shift to the other alternative may not have been completed in time. Thus, when the response alternatives were again available, the most easily accomplished response was emitted. This line of reasoning implies that, although there is some apparent discrimination of the local contingencies of reinforcement (in these cases the strict alternation of reinforcers), this discrimination is by no means perfect. Some reinforcers on the not-just-productive alternative may be miscategorised as reinforcers on the just-productive alternative (Davison & Jenkins, 1985; Davison & Nevin, 1999), thus maintaining responding on the just-productive-alternative when such responses are more easily emitted (as in when the post-reinforcer changeover contingencies favour them).

Despite an apparent initial tendency to stay at the alternative first illuminated after a reinforcer, there was a clear trend for preference in all of the strict alternation conditions to be towards the not-just-reinforced alternative. This was demonstrated in the preference pulses (Figures 5.2), and was more apparent when these preference pulses were plotted with the time taken to make changeover responses removed (Figures 5.4). The probability of staying at the just-reinforced alternative in the strict and random alternation conditions also indicated an effect of the strict alternation contingencies. Across subjects, in the random-alternation conditions, the probability of staying was greater than .5 in 34 of 60 cases (Figure 5.5). In the strict alternation conditions, the probability of staying was greater than .5 in only 17 of 60 cases (Figure 5.6). Thus, although there was some tendency to stay at the alternative that came on immediately after a reinforcer, behaviour was still controlled by the alternation contingencies.

The alternative (just-reinforced versus not-just-reinforced) that was illuminated after a reinforcer determined the degree of preference for the not-just-reinforced alternative in the strict alternation conditions. Preference for this alternative was greatest when it always came on after a reinforcer, was intermediate when it sometimes came on, and was least when it never came on after a reinforcer. This is rather intriguing. Preference was not a simple function of the just-illuminated alternative, but was rather some function of the global contingencies of post-food keylight illumination: Comparing only incidences of the not-just-reinforced-alternative coming on after a food, preference was further towards the not-just-reinforced alternative in Condition 10, which arranged that the not-just-reinforced alternative always came on after a food, than in Condition 5, when the not-just-reinforced alternative came on only half the time. Similarly, comparing only incidences when the just-reinforced alternative came on after the food, preference was more extreme in Condition 5, when this only occurred 50% of the time, compared to either Condition 4 or Condition 12, where this always occurred. This suggests that there was some learning about the local contingencies of reinforcement throughout the condition. The pigeons were not likely responding according to some simple rule such as “always stay at the first-lit alternative” or “always switch from the first-lit alternative”. Indeed, such a rule would be useless in Condition 5. If a pigeon was using such a rule, preference would be less extreme in Condition 5 compared to Conditions 4 or 12. The finding that preference was more extreme in Condition 5 than in Conditions 4 or 12 suggests that the pigeons were not responding according to any rule based on what alternative was illuminated after the last reinforcer; rather, they were responding according to which alternative was responded to prior to the last reinforcer.

The present results confirmed that local preference is consistently towards the alternative that is perceived as more likely to provide the next reinforcer, or more likely to provide a reinforcer sooner. Global contingencies however also had an effect, as demonstrated by the preference trees in the random alternation conditions — preference increased as the number of successive continuations from one alternative increased. This experiment also confirmed results from previous strict-alternation experiments: alternation can be achieved, although with some difficulty and imperfectly. These difficulties in control by strict-alternation contingencies are partly due to the changeover contingencies, and partly due to a carryover of preference from one IRI into the next.

Chapter VI

6.1 Experiment 4

Preference after a response-contingent event in Experiments 2 and 3 was a function of the local left: right food ratio signalled by that event. Preference after a red or green keylight in Experiment 2 was towards the alternative signalled by that keylight as more likely to provide the next food (c.f., Experiment 1 where the information provided by a keylight was redundant with information already present in the global food ratio). Experiment 3 demonstrated that local preference after a food is also a function of the local food ratio in that period. When foods randomly alternated across the left and right alternatives, preference was generally towards the alternative illuminated immediately after the reinforcer. When foods strictly alternated, preference was generally towards the not-just-reinforced alternative (although preference was further towards the not-just-productive alternative when it was available immediately after the reinforcer). Krägeloh et al. (2005) and Davison and Baum (2006; 2010) also found that preference after a response-contingent food or non-food stimulus was a function of the local food ratio after that event.

How does this control by the local contingencies interact with control by temporally extended, global contingencies? Some researchers regard matching, in which the long-term distribution of responses across two or more alternatives is approximately equal to the long-term distribution of reinforcers across those alternatives (Davison & McCarthy, 1988), as arising from contingencies operating at a local level: for instance, every response is allocated to the alternative momentarily more likely to provide a reinforcer (Shimp, 1966, 1976a, 1976b; Silberberg et al., 1978). According to such *momentary maximization* accounts, behaviour is sensitive to only its immediate consequences. Any apparent sensitivity to longer term aggregations of consequences is derivative from this local-level control. Global-level matching can emerge from a number of different local distributions of behaviour however, and responses are often emitted to the alternative momentarily less likely to produce the next reinforcer (Nevin, 1969). For these reasons, Nevin argued that long-term aggregations of behaviour can approximate long-term distributions of reinforcers directly, without needing to go through an intermediate process operating at a local level. Buckner, Green and Myerson (1993) also reported that, although there were behavioural regularities at a local level (visits to an alternative were longer after a reinforcer), these local-level effects were not responsible for long-term matching. Matching (to the prior reinforcer ratio) was maintained during short as well as more extended periods of extinction, despite the absence of any short-term, reinforcer-produced, effects. This suggests, contrary to momentary maximizing, that behaviour is directly sensitive to temporally extended distributions of consequences. A number of other findings also support direct control by temporally extended contingencies, apparently not mediated by control at a more local level (Jones & Davison, 1997; Neuman, Ahearn, & Hineline, 2000; Wanchisen, Tatham, & Hineline, 1988). Other research however, suggests that behaviour can be sensitive to *both* short-term as well as temporally extended contingencies (Williams, 1991a). When adherence to the short-term contingencies is incompatible with adherence to the longer term contingencies, preference is generally intermediate between the two extremes (Hiraoka, 1984; Zeiler, 1987).

Davison and Baum and colleagues have also reported a number of local-level regularities in choice behaviour. Notably, preference immediately after a reinforcer is typically towards the just-productive or momentarily most likely to reinforce alternative (e.g., Davison & Baum, 2002; 2006). Additionally, the log response ratio in successive IRIs visibly shifts towards the source of the prior reinforcers (e.g., Baum & Davison, 2009; Davison & Baum, 2000). Although Landon, Davison and Elliffe (2002; 2003b) and Experiment 1 found these local regularities in steady-state arrangements, there is no evidence to suggest either that these local level effects are responsible for global level matching, or that they are independent of matching at a global level.

Evidence of local-level control was also found in Experiments 2 and 3: although the overall reinforcer ratio was 1:1, preference after a stimulus in Experiment 2 was a function of the local (9:1 or 1:9) reinforcer ratio. In Experiment 3, preference after a food was to the not-just-reinforced alternative in the strict alternation conditions. Williams (1991a) and Silberberg and Williams (1974) both reported that adherence to the global contingencies (global matching) increased as control by (incompatible) local-level contingencies decreased demonstrating that matching at a global level was not dependent on any control by local contingencies.

Experiment 4 further explored questions about the locus of control in choice procedures. Reinforcers appeared in strictly alternating sequences of specified lengths (which varied across conditions). For example, in Condition 3 a sequence of 3 left reinforcers was always followed by a sequence of 3 right reinforcers. After the first and second reinforcers in the sequence, the probability of a continuation was 1.0. After the third reinforcer in the sequence, the probability of a continuation was 0.0. Thus, the overall probability of a continuation in Condition 3 was (1.0+1.0+0.0)/3 = .67. Increasing the sequence length to 5 (Condition 6) also increased the overall probability of a continuation to .80; and decreasing the sequence length to 2 (Condition 11) decreased the overall probability of a continuation to .50. In Condition 2, reinforcers appeared in sequences of 3, but the sequences randomly alternated. In this condition, the overall probability of a continuation was (1.0+1.0+0.5)/3 = .83. In all of these conditions the local probability of a continuation as potentially signalled by the prior reinforcer alternated between extreme levels of 1.0 and 0 (or 0.5 in the case of Condition 2). Will preference be a function of these local probabilities of a continuation? Or will preference be a function of the global probability of a continuation, as found by Krägeloh et al. (2005)? Krägeloh et al. directly manipulated the overall probability of a continuation and found that preference, measured by preference pulses and preference trees, was further towards the just-productive alternative when the probability of a continuation was higher. If choice is also a function of the overall probability of a continuation in Experiment 4, then it will be more extreme at all positions within a sequence of same-alternative reinforcers in conditions with longer sequences. If, however, choice is sensitive to the local probability of a continuation, preference after a discontinuation will be similar whether sequences are 2, 3, 4.5 or 5 reinforcers long. In all of these cases, there is a 1.0 probability of a continuation following a discontinuation.

Perfect adherence to the local contingencies of reinforcement in this experiment required that choice come under the control of the number of preceding same-alternative reinforcers. This may in turn require some numerical competence, or the ability to discriminate the number of preceding same-alternative reinforcers. . Although counting may not be required per se (Davis & Perusse, 1988), different contingencies were operative after the second (or third or fifth) reinforcer in a sequence compared with after the first. There is some rather compelling evidence that the number of experienced biologically-relevant events can exert control over behaviour. Davis and Memmott (1983), Davis and McIntire (1969) and Seligman and Meyer (1970) all reported such evidence in a procedure where each session of food-reinforced responding was interspersed with three unsignaled, unavoidable shocks. Responding was suppressed after the first two shocks to a greater degree than after the third shock, suggesting that the three shocks, taken together, signalled the absence of shock in the remainder of the session. Control by the number of received shocks, as opposed to elapsed time in the session, was supported by the fact that there was no difference in response rates whether all three shocks were presented early in the session, or whether the second shock was presented late in the session: responding was always lower after the first and second shocks than after the third.

Davis and Memmott (1983) noted that although the rats could use the number of received shocks as a discriminative stimulus, they did so only when nothing else signalled shock safety: when the animals could use temporal autocontingencies to predict shock safety, they did so preferentially. In these cases, there was no control by the number of received shocks. Davis and Memmott proposed a "last resort hypothesis" — while rats *can* use the numerical attributes of the situation as a discriminative stimulus, they would rather not. Rats instead prefer to attend to other features of the environment (e.g., the temporal characteristics) and do so more easily.

Conversely, Capaldi and Miller (1988a; 1988b) argued that numerical features of the situation easily acquire discriminative control, even when other aspects (e.g., temporal) are relevant to the discrimination. In a series of runway experiments, Capaldi and Miller arranged a particular sequence of runways ending in food, followed by a runway which did not. For example, the first three runways always contained food and the fourth was always empty. After consuming the third food, the rats ran slower in the final runway, indicating that they anticipated the forthcoming empty runway, and that this discrimination was dependent on the number of consumed foods. Various control conditions were run to eliminate control by temporal features of the experiment (e.g., the time held in each runway was varied). There also exist a number of compelling anecdotal reports suggesting evidence of sophisticated numerical discriminations by nonhuman animals: Egremont and Rothschild (1979) described the ancient Chinese practice of fishing with cormorants. The cormorant dives for fish which it returns to the fisherman. The cormorant is permitted to consume every 8th fish, and apparently refuses to return to the water after retrieving the 8th fish if its neck ring is not loosened, permitting consumption of the 8th fish. As the interval between fish captures is variable, this implies some control by the preceding number of fish retrieved. Egremont and Rothschild provided no details on these birds' training histories, nor did they report any extensive testing of their numerical competence, preventing strong conclusions about such abilities. Nonetheless, this, along with Capaldi and Miller's results, is suggestive of some ability to make sophisticated discriminations based on the number of consumed or unconsumed food items.

The above results imply that rather precise control by the number of consumed same-alternative foods may be readily obtained in Experiment 4. Research on sequential alternation procedures however suggests that control by the number of preceding same-alternative reinforcers is unlikely. As noted in reference to Experiment 3, success in single alternation procedures is obtained with difficulty; preference in one ITI typically carries over to the next. Apparently even more difficult are sequential alternation procedures in which, as in Experiment 4, reinforcers appear in fixed sequences on one alternative which alternate strictly with fixed sequences on another alternative. This task has proven extremely difficult, if not impossible, for nonhuman animals, even primates, and even juvenile humans (Kundey & Rowan, 2009). In a discrete-trials procedure, Williams (1976) found that pigeons showed no evidence of learning the double alternation procedure after 20 sessions. Rather, 5 of the 6 birds perseverated on an alternative until they did not get the reinforcer, then switched to the other alternative. 135 sessions of unsuccessful attempts to obtain double alternation then followed. Eventually, some evidence of control by the double alternation procedure was obtained. However, although double alternation was obtained for the first 4 positions in the sequence (e.g., Response 1, Response 1, Response 2, Response 2), accuracy in longer sequences (e.g., a 5th correct response) was at chance levels. This suggests that rather than learning to "count" same-alternative reinforcers or otherwise base their discrimination on the number of same-alternative reinforcers, success with the first four responses in the sequence was achieved via a long behaviour chain. This apparent inability of non-primate animals to perform in sequential alternation procedures is consistent with research demonstrating that animals abandon a food site containing a fixed number of prey when the IRI exceeds some threshold, and not when the patch had been depleted (Krebs et al., 1974; Roche et al., 1998), contrary to Gibb's (1966) hypothesis of hunting by expectation.

Notably, Lima (1984) found that, although the time spent in an empty patch before abandoning it was inversely related to the number of prey in non-empty patches, all potential food sites in a non-empty patch were sampled, even after that patch was depleted and regardless of the number of foods per non-empty patch. This was the case even when non-empty patches contained food in only 2 of 24 potential sites. The animals were apparently responding according to the overall probability of a patch containing a prey. Once a prey had been found, this overall probability continued to govern behaviour; there was no control by the local probability of the patch containing a prey (which at some point became 0). These results suggest control by the global contingencies (number of prey in a non-empty patch) even in the absence of precise control by the number of consumed foods (number of prey remaining in the current patch).

Kundey & Rowan (2009) hypothesized that double alternation is difficult because the same stimulus (e.g., a left food) sometimes signals a left response, and other times signals a right response. The discriminative stimulus which signals the correct response is not just the preceding single reinforcer, but is the preceding sequence of reinforcers. This adds a further layer of difficulty onto the already difficult single alternation procedure. Thus, choice in Experiment 4 may be a function of the overall probability of a continuation (as in Krägeloh et al. 2005), and there may be no control by the local probability as signalled by the position within the sequence of same-alternative reinforcers (as found by Lima, 1984). The results of Experiments 2 and 3 however, suggest that there may be some control by the local probability of a continuation as signalled by the previous reinforcer. Experiment 3 also confirmed that preference moved further toward the alternative providing reinforcers as the number of preceding same-alternative reinforcers increased. Thus, there is some effect of the number of preceding same-alternative reinforcers. Perhaps discriminative control can be acquired by the preceding sequence of reinforcers rather than simply by the most recent single reinforcer.

6.2 Method

Subjects and Apparatus

The Experiment 3 subjects and apparatus were used in Experiment 4.

Procedure

The conditions of Experiment 4 were interspersed within the conditions of Experiment 3. Two conditions reported in Experiment 3 will also be reported here (although the analyses will differ). Reinforcers were arranged in sequences which either randomly or strictly alternated. This was varied across conditions as was the length of each sequence. Condition 13 replicated Krägeloh et al.'s (2005) procedure. In that condition, the local probability of a reinforcer was .83, the same as the overall probability of a continuation in Condition 2 (in which sequences of 3 randomly alternated), and very similar to the overall probability in Condition 6 (when sequences of 5 strictly alternated). Condition 14 investigated what effect (if any) there was of the post-food changeover contingencies. In Condition 14, the not-just-reinforced alternative became available after a food (as in some conditions of Experiment 3). Table 6.1 presents the sequence of conditions.

Table 6.1

Sequence of conditions in Experiment 4 along with the length of sequences in that condition, whether the sequences strictly or randomly alternated, and the overall probability of a continuation reinforcer in that condition.

Cond.	Sequence length	Strict vs. random alternation	Overall *p*. of a Continuation
1	1	Random	.5
9	1	Random	.5 (Rep of Con 1)
2	3	Random	.83
3	3	Strict	.67
6	5	Strict	.80
11	2	Strict	.5
13			.83
14	3	Strict	.67
15	3	Strict	.67 (Rep of Con 3)

The data from Conditions 1 and 9 were also analyzed in Experiment 3. Condition 9 was a replication of Condition 1 and Condition 15 was a replication of Condition 3. Condition 14 was exactly the same as Conditions 3 and 15 except that the not-just-reinforced alternative came on after every food in Condition 14, whereas the just-reinforced alternative came on after every food in Conditions 3 and 15. Other procedural details, including session and condition run-times as well as reinforcer-scheduling and changeover contingencies were as described in Experiment 3.

6.3 Results

Figure 6.1 presents the preference pulses from conditions with an overall continuation probability of .5, or an average sequence length of 2. The data from two of these conditions (Condition 1 and its replication Condition 9) were also presented in Experiment 3. The analyses in Figure 6.1, however, presents local preference as a function of the location of not just the most recent food, but also as a function of the location of the food prior to the most recent one. Condition 11 arranged strict alternation of sequences of 2 same-alternative reinforcers. Condition 11 differed from Conditions 1 and 9 in that the local probability of a continuation reinforcer was either 1.0 or 0 in Condition 11, while it was always .5 in Conditions 1 and 9. Appendix Figures J1 to J6 present this figure for the individual subjects. Consistent with Experiments 2 and 3, data from the last 20 sessions of each 65 session condition were taken for analysis. For each individual, the log (left/right) response ratio was calculated in each 2-s time bin after left and right foods preceded by either a same- or an other-alternative food. A value of ± 3.5 was used if preference was exclusive to one or the other alternative. If a time-bin contained fewer than 20 responses in total, no value was calculated for that time bin. Only if 2 or more individuals had valid data points in a time bin was a group mean value calculated.

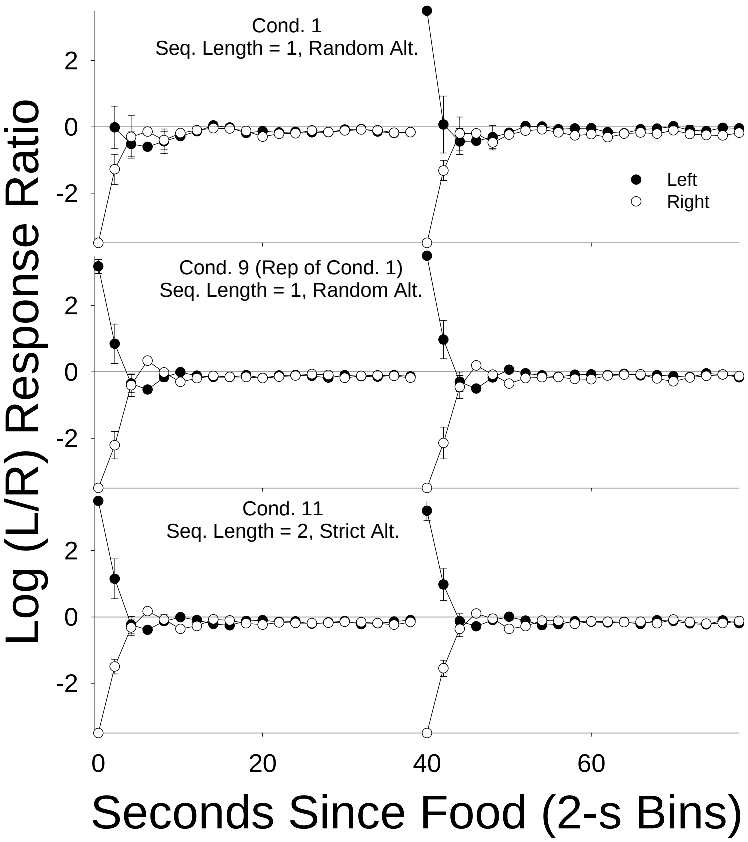


Figure 6.1 Group mean log (left/right) response ratio in the first 40 s following a reinforcer in conditions with a .5 probability of a same-alternative reinforcer. In Conditions 1 and 9 reinforcers randomly alternated. In Condition 11, sequences of two reinforcers from one alternative strictly alternated with sequences of two reinforcers from the other alternative. The leftmost preference pulses in each plot show preference after a reinforcer which differed from the one immediately prior while the rightmost preference pulses show preference after a reinforcer whose location was the same as the last reinforcer. Error bars represent one standard error and are placed at representative data points increasing in $\log_2$ units.

In the three panels of Figure 6.1, the preference pulse after a left food preceded by a left food was very similar to the preference pulse after a left food preceded by a right food. The same was true for local preference after a right food preceded by either a left or a right food. There was no detectable effect of whether the most recent reinforcer was a continuation or a discontinuation in any of the conditions where the overall probability of a continuation was .5 in either the group, or the individual-subject (Appendix Figures J1-J6) preference pulses. Although

Experiment 3 reported that there were effects of reinforcer sequences in Conditions 1 and 9, these effects were apparently not detectable in sequences of only 2 same-alternative reinforcers. In Condition 11 the probability of a continuation reinforcer was 1.0 after a discontinuation, and was 0 after a continuation. Despite this, the preference pulses were very similar after a continuation and after a discontinuation. Additionally, preference pulses in Condition 11 did not differ from preference pulses in Conditions 1 and 9, suggesting that, although the probability of a continuation oscillated in Condition 11 between the two extreme values of 0 and 1.0, preference was controlled by the average of these: a probability of a continuation of .5.

Figure 6.2 and Appendix Figures J7 to J12 present local preference when reinforcers appeared in strictly alternating sequences of 3. In all of these conditions, the probability of a continuation reinforcer once again depended on the prior reinforcers. After a discontinuation and after the first continuation, the probability of a continuation was 1.0. After the second continuation, the probability of a continuation was 0. The overall probability of a continuation in these conditions was .67. In Conditions 3 and 15, the just-productive alternative always came on after a reinforcer (as in all other conditions of Experiment 4). In Condition 14 the not-just-reinforced alternative always came on.

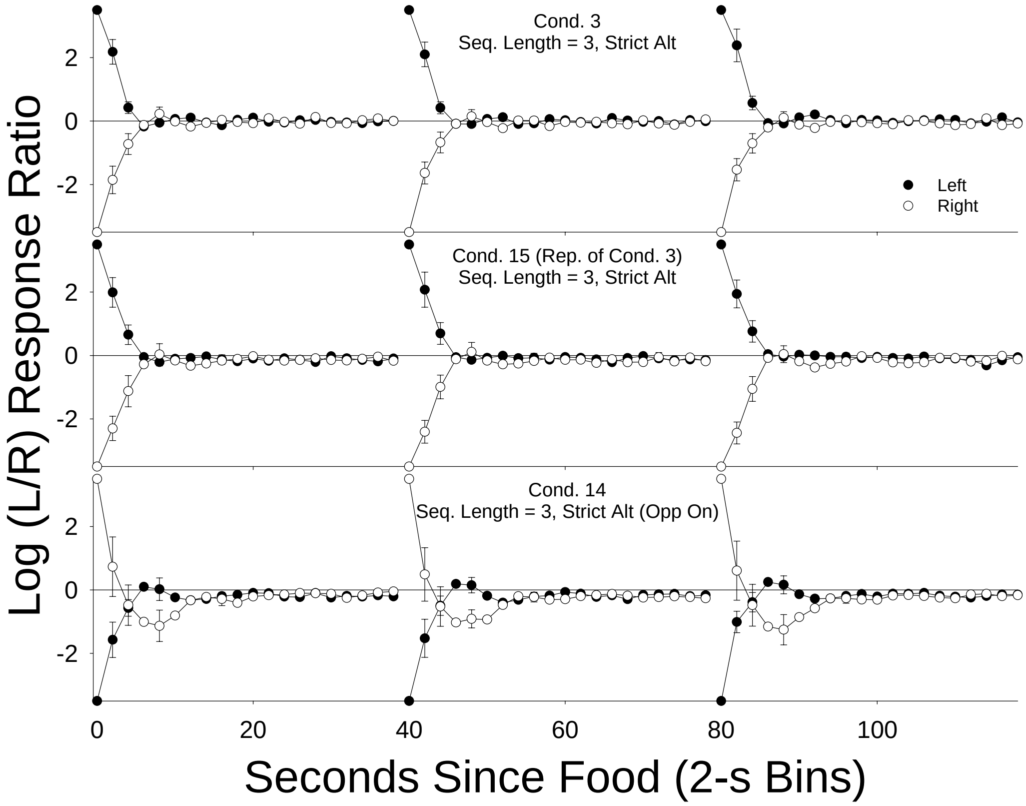


Figure 6.2 Group mean log (left/right) response ratio in the first 40 s following a reinforcer in conditions with a .67 probability of a same-alternative reinforcer. Sequences of 3 reinforcers from one alternative strictly alternated with sequences of 3 reinforcers from the other alternative. In Condition 3 (and its replication Condition 15), the just-reinforced alternative always came on after a reinforcer. In Condition 14, the not-just-reinforced alternative always came on. Preference pulses are plotted as a function of where in the sequence of continuations the prior reinforcer fell (first, second or third). Error bars represent one standard error and are placed at representative data points increasing in $\log_2$ units.

Figure 6.2 (and the individual-subject plots; Appendix Figures J7-J12) suggests no effect of either the preceding number of continuations, or the probability of a further continuation. When sequences of three reinforcers strictly alternated, preference was equally towards the just-productive alternative at all sequential positions. Preference after a food in Condition 14 (when the not-just-productive alternative was illuminated after a food) was initially towards the not-just-productive alternative before briefly turning to the just-productive alternative, finally settling at relative indifference. This mirrors preference in Experiment 3 when the just-productive alternative was available immediately after a food and reinforcers strictly alternated. In those conditions, preference was also first towards the alternative that was illuminated after the reinforcer before moving to the other alternative (the alternative more likely to provide the next reinforcer). Condition 14 demonstrates that this pattern of preference was not a function of the number of preceding same-alternative reinforcers.

The lack of an effect of the local probability of a continuation (0 or 1.0) demonstrated in Figures 6.1 and 6.2 may have been expected given previous demonstrations that sequential alternation is unlikely. The absence of any effect of the number of preceding continuations is perhaps somewhat more unexpected. Typically, preference pulses become more extreme and remain further from indifference as the number of preceding same-alternative reinforcers increases. This was even demonstrated with these pigeons in Experiment 3. Were the preference-increasing effects of reinforcer sequences counteracted by the preference-attenuating effects of the forthcoming sequence of other-alternative reinforcers? Before this conclusion can be accepted, it must be determined whether 2- or 3-reinforcer sequences are simply too short to demonstrate typical sequence effects. Figure 6.3 presents preference pulses from conditions with longer sequences. In Condition 6, sequences of 5 reinforcers from one alternative strictly alternated with sequences of 5 from the other. In Condition 3, sequences of 3 reinforcers randomly alternated, and Condition 13 replicated Krägeloh et al.'s (2005) procedure: a reinforcer was delivered for a response to the just-productive alternative with a probability of .83 (regardless of the number of preceding continuations). Appendix Figures J13 to J18 present the individual subject preference pulses.

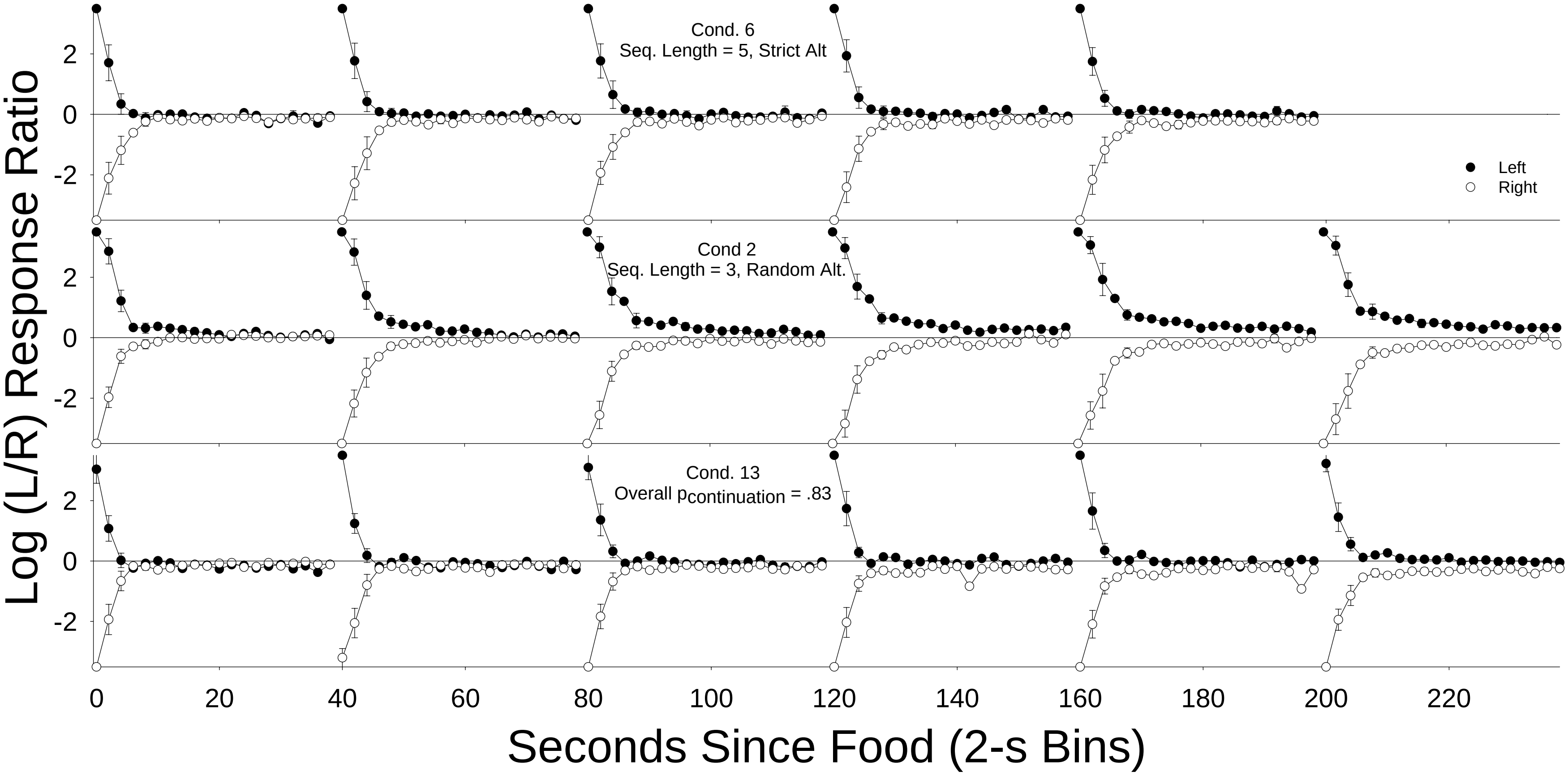


Figure 6.3 Group mean log (L/R) response ratio after in the first 40 s after a reinforcer when the overall probability of a continuation was about .8. In Condition 6 (top panel), 5-reinforcer sequences strictly alternated (overall p_{same} = .8). In Condition 2 (middle panel), 3-reinforcer sequences randomly alternated (overall p_{same} = .83). In Condition 13 (bottom panel), the probability of a continuation after any reinforcer was .83. The number of preceding same-alternative reinforcers increases from left to right in each plot. In the middle panel, the preceding sequence of 3 reinforcers was on the other alternative in the 6 left-most plots, and was on the same alternative in the 6 right-most plots. Error bars show ±1 standard error and are at representative data points increasing in $\log_2$ units.

There was some tendency for the preference pulses in Figure 6.3 (as well as those in Appendix Figures J13 to J18) to move apart as the number of preceding same-alternative reinforcers increased. This trend was present in all three conditions, but appeared greatest in Condition 2 (random alternation of 3 reinforcer sequences). However, this may be due to a small point of difference between how the Condition 2 data were divided for plotting in Figure 6.3 versus how the Condition 6 and Condition 13 data were divided. In the top and bottom panels of Figure 6.3 the 4th preference pulses from the left depict preference after 4 reinforcers on the same alternative preceded by either nothing (the start of the session), or a reinforcer on the other alternative. The Condition 2 preference pulses on the other hand (middle panel), depict preference after *at least* 4 reinforcers on the same alternative. If a reinforcer had been delivered prior to the fourth one, it was equally likely to be on the left or the right key. Thus, in about half of the cases, the sequences are actually longer than depicted in (the second half of) Condition 2. Nevertheless, Figure 6.3 still demonstrates that although no such effect may be apparent in Figures 6.1 and 6.2, sequences of same-alternative reinforcers did have their typical preference-increasing effects. This occurred whether long sequences of same-alternative reinforcers strictly alternated, slightly shorter sequences of same-alternative reinforcers randomly alternated, or the sequences were long and variance in sequence length was maximal. The cumulative effects of successive same-alternative reinforcers are apparently rather small, and are only evident in relatively long sequences.

The effects of reinforcer sequences may be more easily seen in preference trees. Figure 6.4 presents the (group mean) continuation trees for conditions where the average sequence length was 2 (the conditions presented in Figure 6.1).

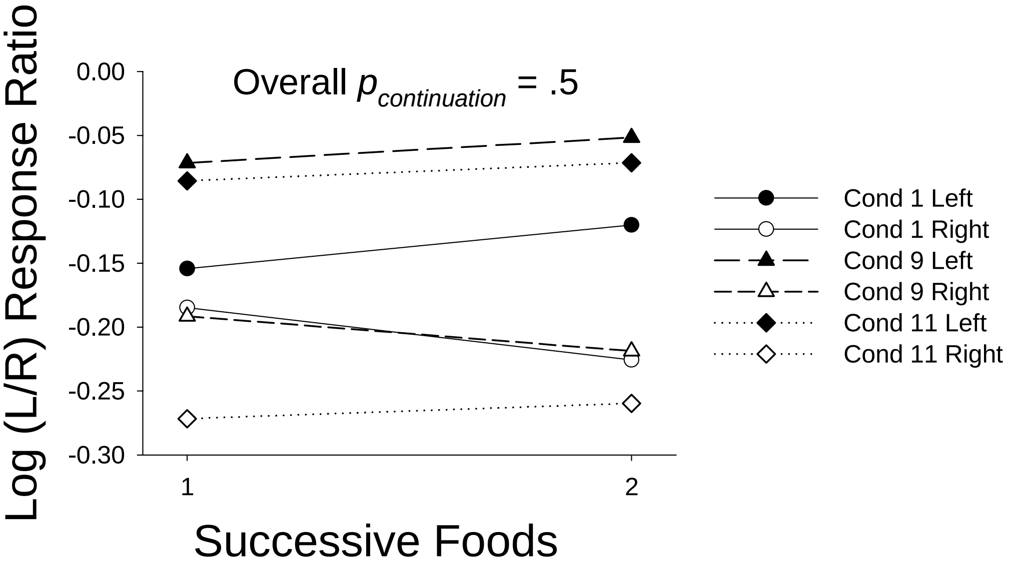


Figure 6.4 Group mean log (L/R) response ratio as a function of the number of successive same-alternative foods in Conditions 1 and 9 which arranged random alternation of reinforcers and Condition 11 which arranged strict alternation of 2-reinforcer sequences.

Wilcoxon matched-pairs signed-ranks tests were conducted on the individual-subject data comparing the log response ratio after the second successive reinforcer from an alternative with the log response ratio after the first reinforcer from that alternative. Two tests were conducted for each condition, one for reinforcers from the left, and one for reinforcers from the right. Preference was always significantly further towards the just-productive alternative after the second reinforcer than after the first reinforcer in Condition 1, for right reinforcers only in Condition 9 and for neither left nor right reinforcers in Condition 11 ($\alpha = .05$). Thus, while there is some evidence that preference shifted towards the alternative providing reinforcers after two reinforcers when the reinforcers randomly alternated, this trend was less evident when reinforcers appeared in fixed sequences of two.

A trend for preference to continually increase as successive continuation reinforcers are delivered may not be apparent after only two reinforcers. Nonparametric trend tests were conducted on preference following the first 8 continuation reinforcers from the same alternative in Conditions 1 and 9. Significant trends were found ($\alpha = .05$) for preference after left and right foods in both Conditions 1 and 9. Such a trend test could not be conducted in Condition 11 because sequences were only ever two long. However, the Wilcoxon tests above indicated no apparent trend for preference after the second reinforcer to be more extreme than preference after the first reinforcer. Figure 6.5 presents preference after each of the 3 reinforcers in the 3-reinforcer sequences arranged in Conditions 3, 15 and 14.

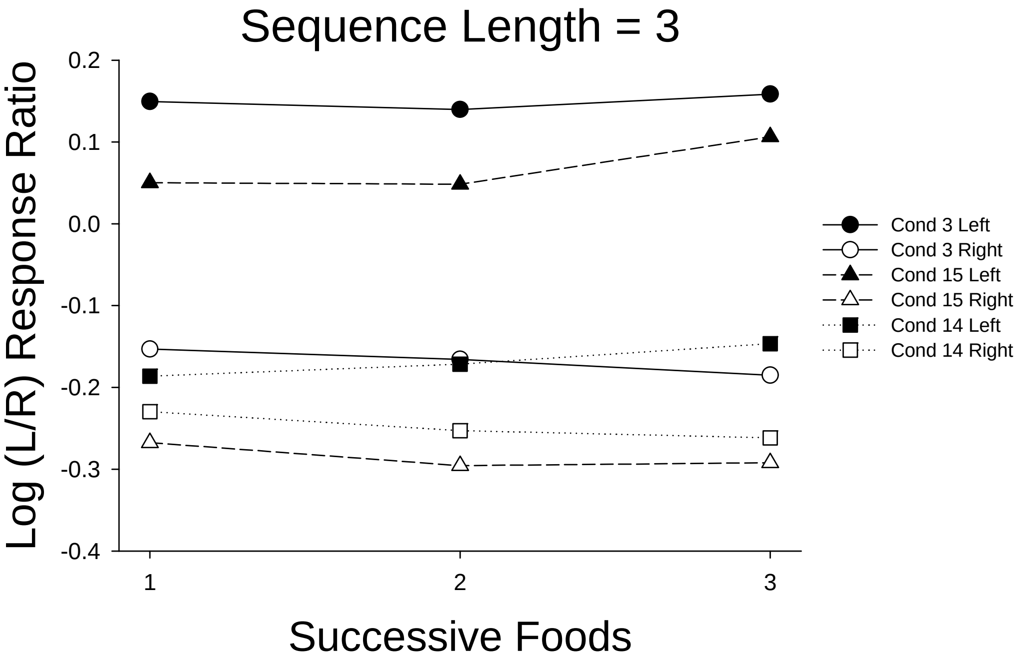


Figure 6.5 Group mean log (L/R) response ratio as a function of the number of successive same-alternative foods in conditions arranging strict alternation of 3-reinforcer sequences. The just-productive alternative always came on after a reinforcer in Conditions 3 and 15 and the not-just-reinforced alternative always came on after a reinforcer in Condition 14.

Only one of the six trend tests conducted (2 for each condition: one for preference after left reinforcers and the other for preference after right reinforcers) revealed a significant trend for preference to move further towards the alternative providing those reinforcers (left reinforcers in Condition 14). The power of a trend test to detect a trend (if present) is low with only three data points and six subjects. When similar trend tests was conducted using only the first three continuation reinforcers in Conditions 1 and 9 (which had revealed significant trends when 8-reinforcer sequences were evaluated), only 1 of the 4 tests indicated a significant trend (preference after left reinforcers in Condition 1). As perhaps a less conservative test, Wilcoxon matched-pairs signed ranks tests compared preference after the first and after the third reinforcer in a sequence. Preference was always significantly further towards the right after the 3rd consecutive right food than after the first right food. Preference was further towards the left after the third consecutive left food than after the first left food in Condition 15 only. Thus, there was some (perhaps weak) evidence of greater preference for the just-productive alternative after the 3rd same-alternative reinforcer compared to preference after the first reinforcer in that sequence. Figure 6.6 presents the preference trees in conditions with an approximately .8 probability of a same-alternative reinforcer (Conditions 13, 2 and 6).

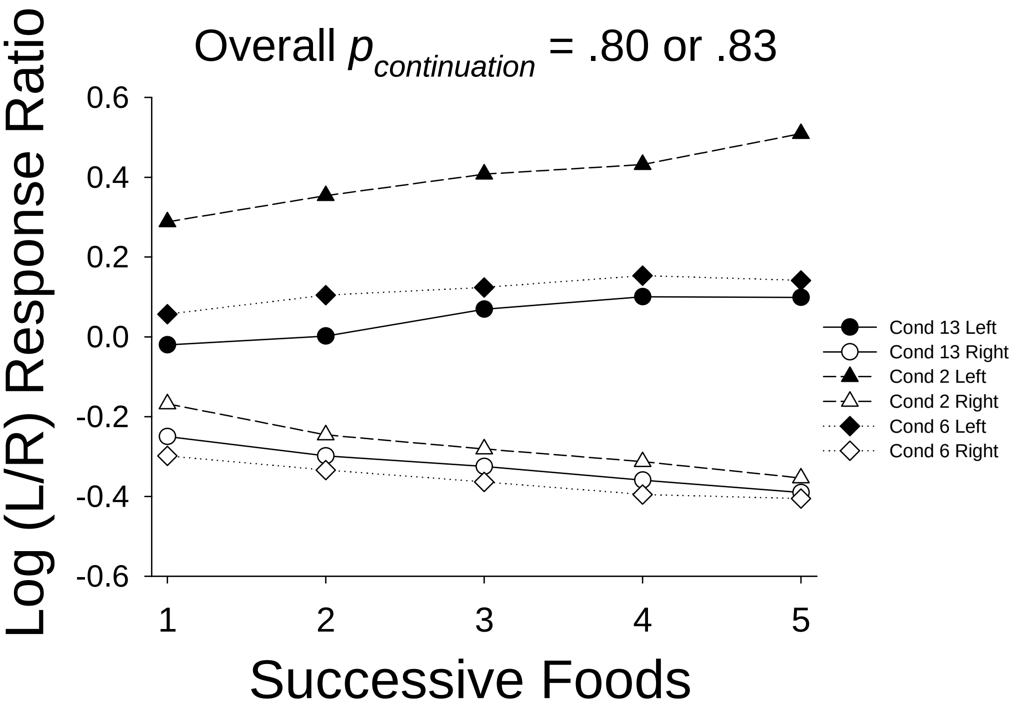


Figure 6.6 Group mean log (L/R) response ratio as a function of the number of successive same-alternative foods in conditions arranging an overall probability of a continuation of about .8. In Condition 13 the overall probability of a continuation was .83. In Condition 2 3-reinforcer sequences randomly alternated, and in Condition 6 5-reinforcer sequences strictly alternated.

Nonparametric trend tests conducted on the individual-subject data indicated that, in all three of these conditions with an overall probability of a continuation of about .8, there was a significant trend for preference to move further towards the alternative providing the reinforcers as successive same-alternative reinforcers were delivered. This was the case whether sequences of 5 strictly alternated, whether sequences of 3 randomly alternated or whether there was a constant probability of a continuation.

Krägeloh et al. (2005) reported that the overall probability of a continuation influenced preference pulses and preference trees. This was also the case in Experiment 4. There appeared to be no control by the local probability of a continuation. Preference was the same whether this local probability was 1.0, .83,

.50, or 0.0. The probability of a continuation was 1.0 after a discontinuation in Conditions 11 (sequence length = 2), 3 and 15 (sequence length = 3), 6 (sequence length = 5) and 2 (sequence length = 3, randomly alternating making the average sequence length = 4.5). The overall probability of a continuation varied across these conditions. Was there any evidence that preference after a discontinuation was controlled by the local probability of a continuation (1.0)? Figure 6.7 presents the discontinuation preference pulses from these conditions (also presented in Figures 6.1-6.3 but presented here together to aid in comparison).

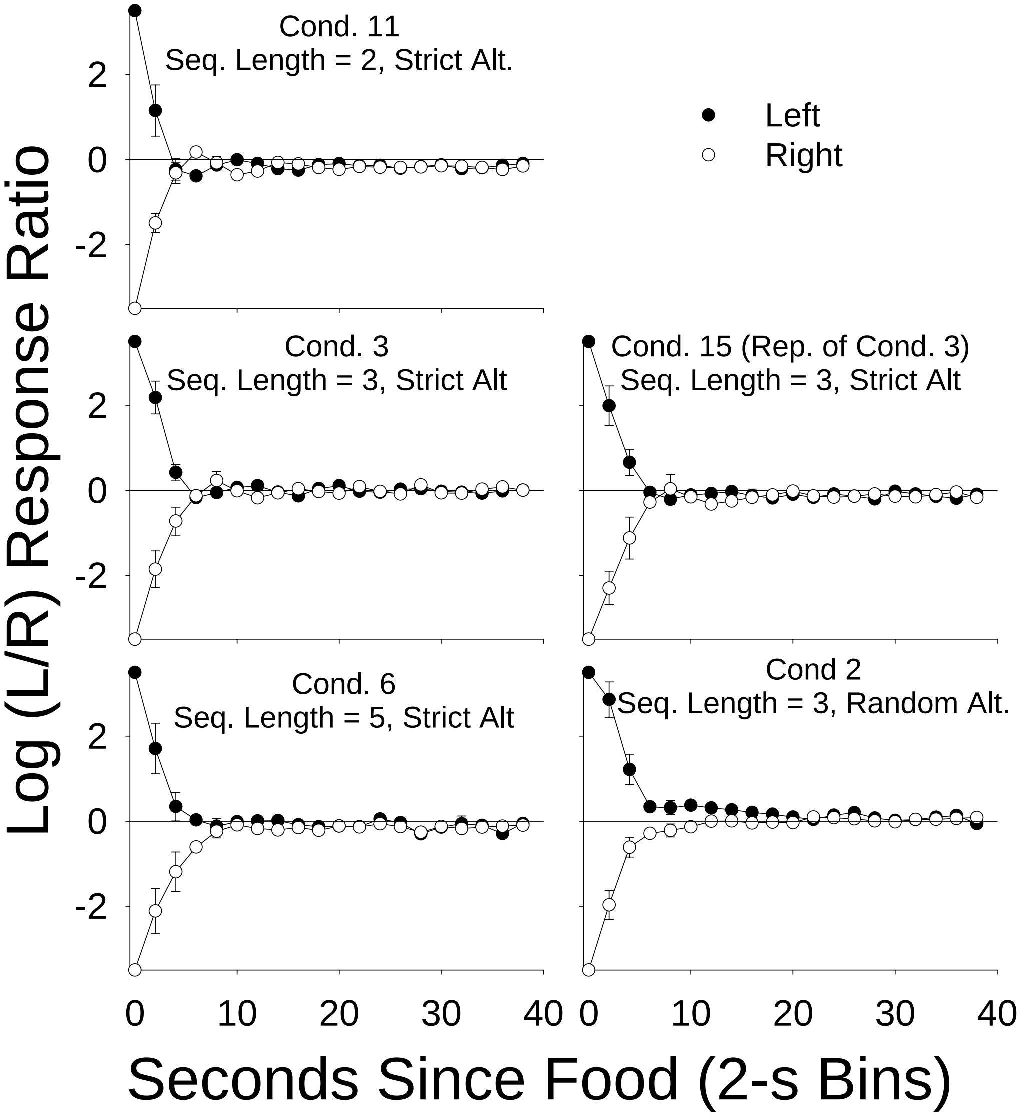


Figure 6.7 Group mean log (L/R) response ratio in the first 40 s after a discontinuation reinforcer (a reinforcer from the alternative that did not provide the immediately prior reinforcers) in all conditions where reinforcers appeared in sequences.

The local probability of a reinforcer on the just-reinforced alternative was 1.0 in all the plots in Figure 6.7. Despite this, preference was clearly further towards the just-productive alternative in conditions with longer sequences, or with a higher overall probability of a discontinuation. This again indicates no control by the strict alternation of fixed sequences of reinforcers: preference was a function of the overall probability of a continuation and was not sensitive to the fact that, in all of these conditions, a discontinuation signalled, with $p = 1.0$, that the next reinforcer would be on the just productive alternative.

The probability of staying at the just-productive alternative may also reveal some effect of the sequence of previous reinforcers. Table 6.2 presents the probability of staying at the just-productive alternative after left and right foods in each condition of Experiment 4.

Table 6.2

Probability of making the first response to the just-productive alternative after left and right foods for each individual pigeon in each condition of Experiment 4.

	Pigeon 141		Pigeon 142		Pigeon 143		Pigeon 144		Pigeon 145		Pigeon 146	
Cond.	L	R	L	R	L	R	L	R	L	R	L	R
1	.16	.92	.65	.74	.99	.67	.80	.91	.22	.85	.02	.98
9	.99	.96	.99	.73	.56	.98	.52	.93	.93	.92	.44	1.00
11	1.00	.98	1.00	.92	.72	1.00	.37	.28	.99	1.00	.75	.97
3	1.00	.98	.99	.93	.95	.81	.89	.97	1.00	.96	.97	1.00
15	1.00	1.00	1.00	.92	.94	1.00	.92	.99	1.00	.98	.97	.98
14	.53	.29	.07	.02	.14	.22	.56	1.00	.16	.16	.07	.11
6	1.00	.99	.99	.66	.97	.99	.75	.97	1.00	.96	.94	1.00
2	1.00	.97	.96	.98	.98	.99	.99	.99	1.00	1.00	.99	.89
13	1.00	1.00	1.00	.92	.83	.97	.46	.97	.95	.96	.89	.86

The probability of staying at the just-productive alternative after a reinforcer was universally high in all conditions except for Condition 14. In Condition 14, the not-just-productive alternative came on after a reinforcer, so this apparent trend to start responding to the not-just-productive alternative, may be better understood as a tendency to respond to the first-illuminated alternative. A Friedman nonparametric analysis of variance was conducted on all of the probabilities in Table 6.2 (except those from Condition 14). There were no differences in the probability of staying across conditions. Thus, the pigeons were always highly likely to stay at the just-productive alternative, regardless of the average sequence length. Figure 6.8 investigates whether the probability of staying after a reinforcer changed as a function of the number of same-alternative reinforcers in any condition. In Figure 6.8, the log ratio of first visits to the just-reinforced alternative over first visits to the not-just-reinforced alternative is plotted as a function of the number of preceding same-alternative reinforcers. Condition 14 is not included in Figure 6.8 as the probability of staying at the just-reinforced alternative was universally low in Condition 14. The individual subject plots were first created (Appendix Figures J19-J24). A valid data point had to exist for at least 2 individual subjects for a data point to be plotted for the group (Figure 6.8). If the probability of staying was ever 1.0, a log (stay/switch) value of 2.5 was plotted. If there were fewer than 10 visits in a bin, no log (stay/switch) value was calculated.

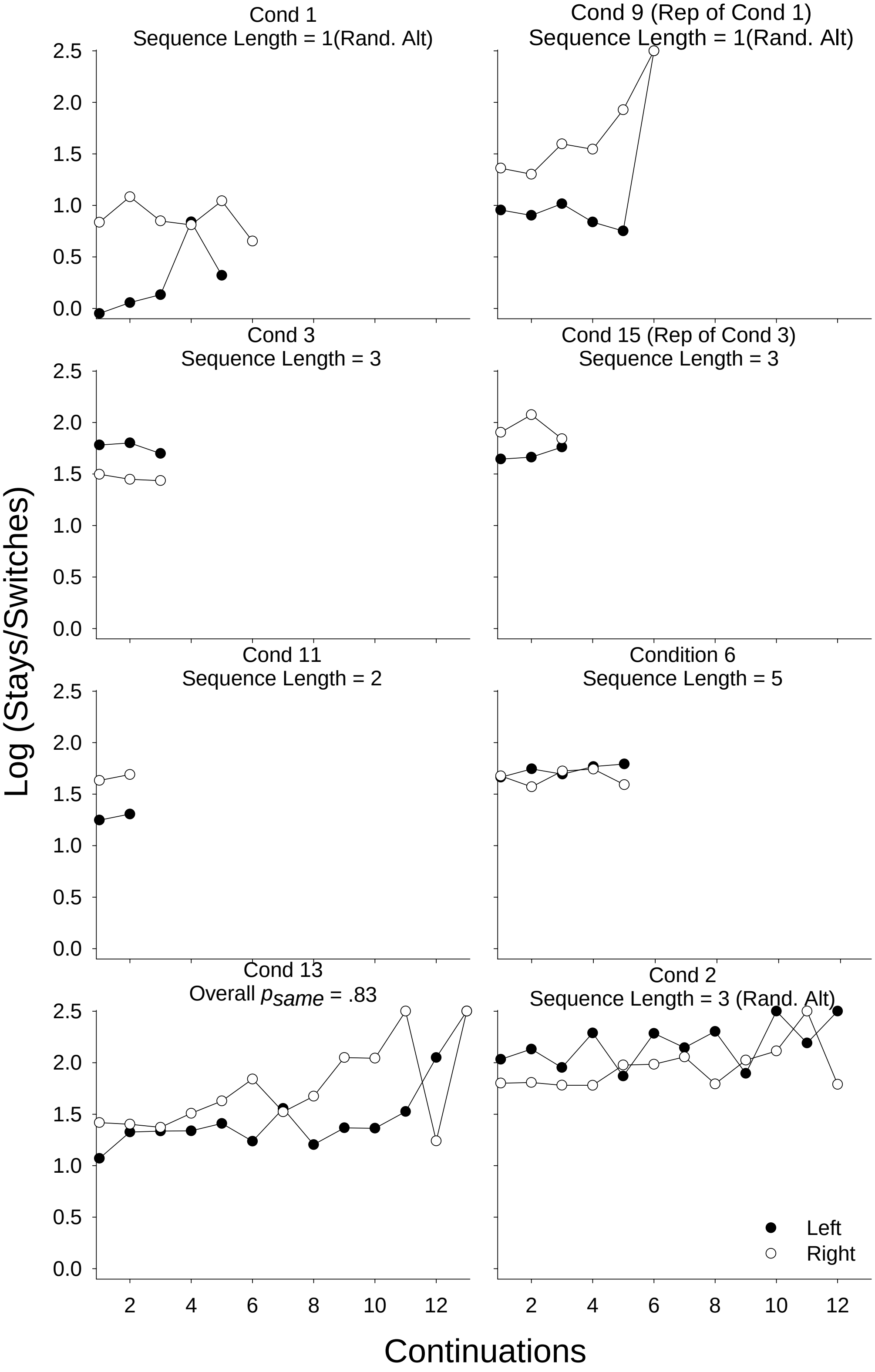


Figure 6.8 Group mean log (first visits to the just-reinforced alternative/first visits to the not-just reinforced alternative) as a function of the number of preceding same-alternative foods delivered without an intervening food from the other alternative.

There was a significant trend for the probability of staying at the just-productive alternative to increase as the number of preceding same-alternative reinforcers increased after both left and right foods in Condition 2. The probability of staying increased as a function of successive right foods only (not after successive left foods) in Condition 6 (strict alternation of sequences of 5), Condition 13 (overall probability of a continuation = .83) and Condition 3 (strict alternation of 3-reinforer sequenes). As the probability of staying was universally high (Table 6.2), there may have been a ceiling effect. Thus, this measure may not be sensitive to detecting any effects of sequential position or overall probability of a continuation. Nonetheless, it did detect some tendency for the probability of staying to increase as successive same-alternative reinforcers were delivered. Figure 6.9 presents the probability of staying after a discontinuation reinforcer (measured as log first stays/first switches) as a function of the number of preceding (other alternative) continuations in all conditions where sequences did not strictly alternate. The individual subject plots are presented in Appendix Figures J25-J30.

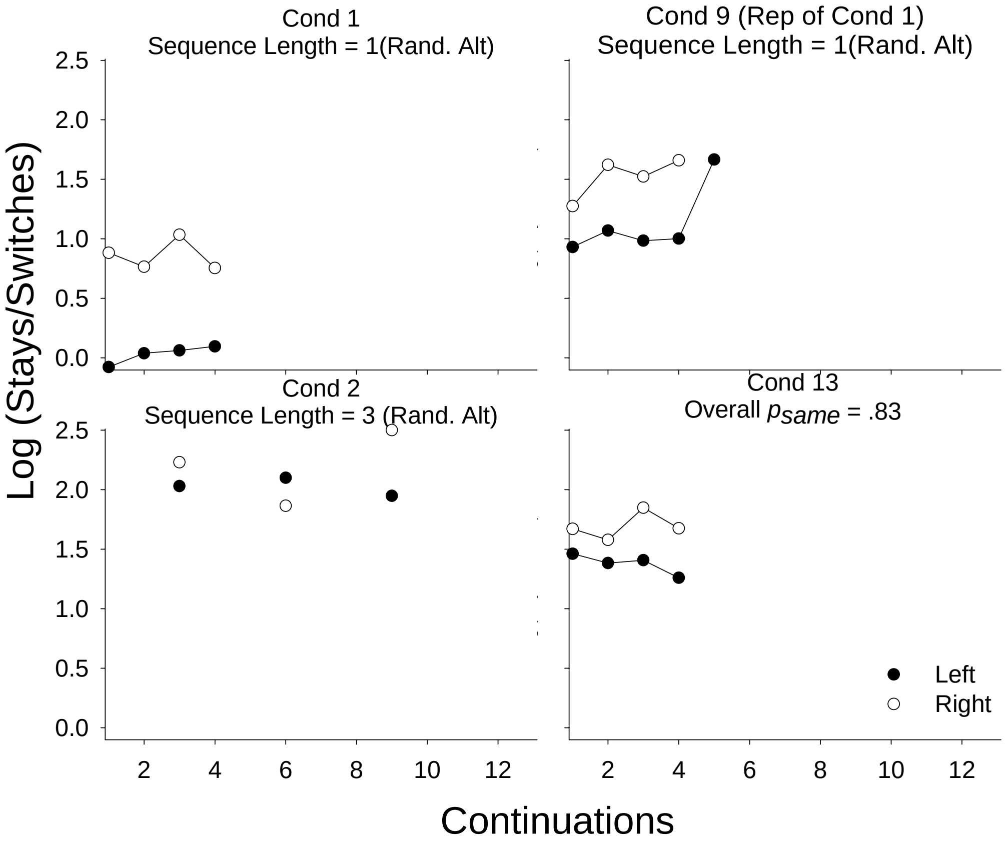


Figure 6.9 Group mean log (first visits to the just-reinforced alternative/first visits to the not-just reinforced alternative) after a discontinuation reinforcer as a function of the number of preceding (other alternative) continuations.

There appeared to be some tendency for the probability of staying at the just-productive alternative to (sometimes) decrease as the number of preceding reinforcers on the other alternative increased. Significant trends were found in Conditions 2 and 13 for both left and right foods, and after right foods only in Condition 9. There were no significant trends in Condition 1. This analysis again suggests some small effect of the preceding reinforcers: the probability of staying after a discontinuation sometimes decreased as the preceding sequence of other alternative reinforcers increased. Figure 6.10 presents the (group mean) pecks in the first visit to the just-reinforced alternative (visit duration; Baum & Davison, 2004) as a function of the number preceding continuations on that alternative. The individual subject plots are presented in Appendix Figures J31-J36.

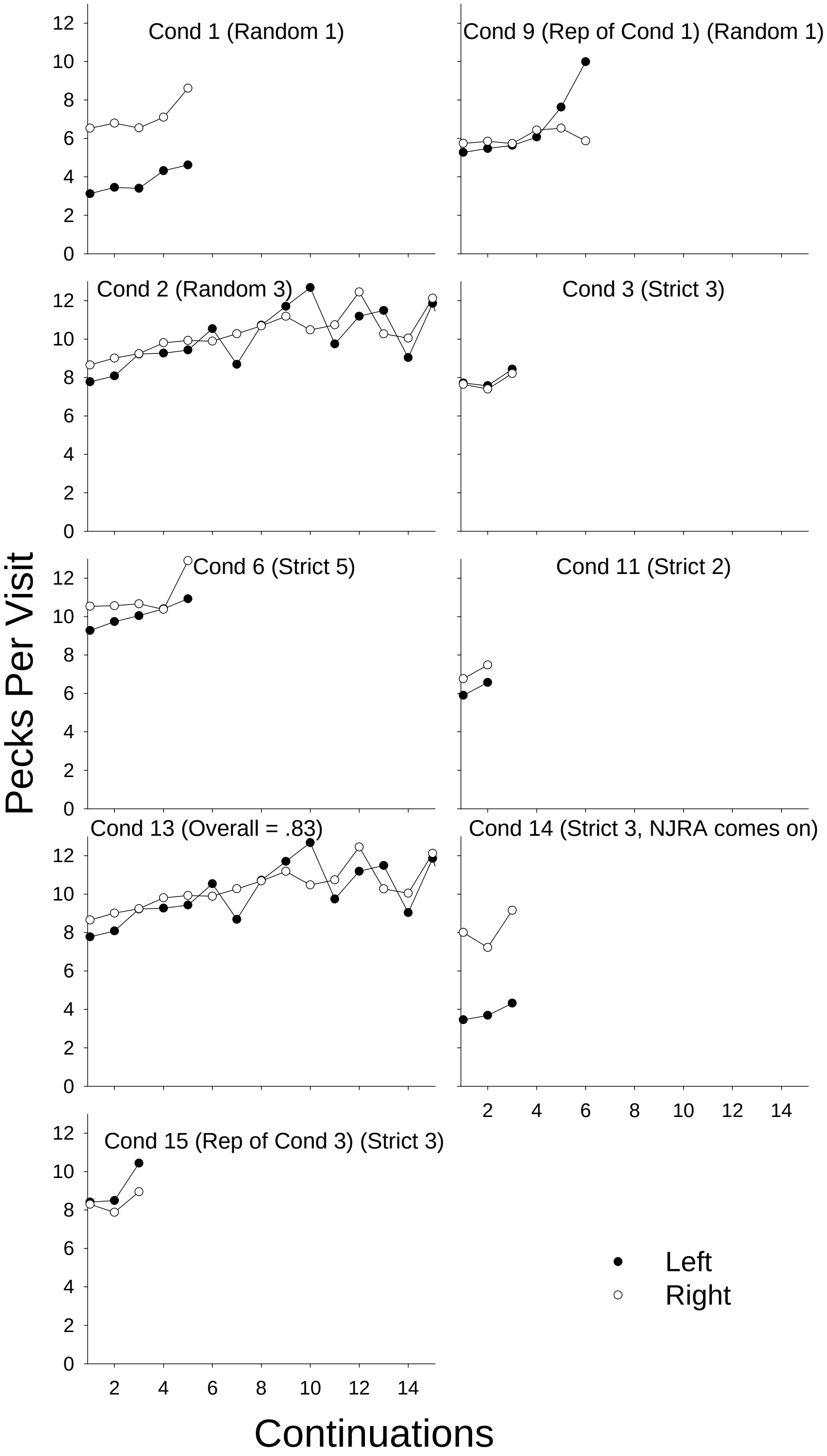


Figure 6.10 Group mean pecks in the first visit to the just-productive alternative as a function of the number of preceding same-alternative foods delivered without an intervening food from the other alternative.

The number of pecks in the first visit to the just-productive alternative clearly increased as the number of preceding same alternative reinforcers increased. Of the 16 trend tests conducted on the data of Figure 6.10 (no trend test was conducted on the Condition 11 data), significant trends were found in 10 tests. There were significant trends for the first (stay) visit to the left to increase in Conditions 2, 6, 13, 9, 3 and 15 and for the first (stay) visit after a right food to increase in Conditions 2, 6, 13 and 14.

If the pigeons could discriminate the start of a sequence of same-alternative reinforcers, the probability of staying and/or the size of the first visit to the just-productive alternative after a discontinuation would be universally high in all conditions where reinforcers appeared in fixed sequences (when the probability of a continuation was 1.0). Nonparametric analyses of variance were conducted to compare the size of the first visit after a discontinuation reinforcer across conditions. Separate ANOVAs were conducted for visits after left and right foods in order to rule out any effects due to position. The ANOVA for responding after a right discontinuation was not significant (p = .47), indicating that the size of the first visit to the right after a right discontinuation was uniform across all conditions (whether reinforcers appeared in fixed sequences or not). Although the ANOVA was significant for first visits to the left after a left food (p = .01), posthoc tests revealed no significant pairwise comparisons. Thus, there were no differences, in the size of the first visit to the just-reinforced alternative. This again confirms that there was no effect of the local probability of a continuation.

6.4 Discussion

The average reinforcer sequence length had an effect on the preference pulses (Figures 6.1-6.3), the probability of staying after a reinforcer (Figures 6.8-6.9), and the size of the first post-food visit (Figure 6.10). The variation in sequence length however had no detectable effect. Thus Krägeloh et al.'s (2005) findings were replicated: both the probability of staying at the just-reinforced alternative and local preference after a reinforcer were controlled by the overall probability of a continuation. There was no effect of changing the local probability of a continuation from 1.0 to 0 within the same condition.

In one sense, these results were expected: previous research has demonstrated that while alternation across response locations is possible (though difficult), sequential alternation is all but impossible for nonprimates. When alternation after multiple same-alternative reinforcers is achieved, it has usually been after extensive training or complicated correction procedures (Kundey & Rowan, 2009) and, even then, evidence suggests that the alternation is achieved via a complex behaviour chain rather than being based on the number of preceding same-alternative reinforcers (Williams, 1976). Additionally, both Krebs et al. (1974) and Roche et al. (1998) reported that even when patches contained a fixed number of prey, animals never learned to abandon a patch after depleting it; instead, the animals abandoned the patch when the time since the last food exceeded some threshold. Lima (1984) also found that although the number of prey per patch had some influence (animals sampled more potential food sites in empty patches when non-empty patches contained fewer foods), they sampled all potential food sites in a non-empty patch, even after obtaining all of the foods, and even when nonempty patches contained only two foods. Similarly, preference in Experiment 4 was universally towards the just-reinforced alternative. The pigeons continued to return to the just-productive alternative even when the probability of a further reinforcer on that alternative was 0. In both Lima's experiment and the present experiment, there was apparent control by the average number of foods at a particular location, but no control by the number of foods thus far consumed.

In Experiments 2 and 3 preference after a response-contingent event (food in the case of Experiment 3 and a keylight illumination in the case of Experiment 2) varied as a function of the local left: right food ratio in that post-event period. Rather than using the location of the last single food as a discriminative stimulus (as in Experiment 3) the sequential alternation conditions of Experiment 4 required that the location of the last 2, 3 or 5 reinforcers be used as a discriminative stimulus. While reinforcers further back than the most recent do have effects on behaviour (Baum & Davison, 2009; Landon et al., 2002; Experiment 3), Experiment 4 demonstrated the difficulty in these earlier reinforces acquiring *discriminative* control.

Davis and Memmott (1983) agued that while discriminative control by numerical features of the environment was possible, it was not likely if any other aspect of the environment could also acquire discriminative control. There were no other potential discriminative stimuli to signal the local reinforcer ratio in Experiment 4: temporal characteristics of the inter-reinforcer intervals could not signal the local reinforcer ratio as a constant-probability VI schedule arranged all reinforcers. It might thus be argued that if the pigeons *could* make discriminations based on the number of preceding same-alternative reinforcers, they would have here. The lack of any discriminative control by reinforcer sequences may be at least partly be due to the fact that the penalty for failing to discriminate sequential position was small: reinforcers were dependently arranged and changeovers were relatively easily accomplished. This procedure may not therefore be the ideal preparation for investigating an ability which may only emerge with difficultly. However, adherence to the local contingencies of reinforcement was found in similar conditions in Experiments 2 and 3, suggesting that although the penalty for not discriminating the local reinforcer ratios may be small, such discrimination is still possible. Capaldi and Miller (1988a) argued that numerical discriminations are easily obtained. Even when not required by the contingencies of reinforcement, control by the numerical features of the environment is still often obtained. It should be noted that while there are some reports of numerical discriminations by pigeons (e.g., Hirai & Jitsumori, 2009), most evidence of numerical discrimination in nonhuman animals has been obtained with rats. A species difference may thus account for the present failure to find any evidence of control by the number of same-alternative reinforcers.

The numerical discrimination tasks with which nonhuman animals have had the most success require discrimination of the *total* number of foods consumed in some time period (e.g., the current session). Sequential alternation tasks, including the present one, require discrimination of the number of foods *from a particular alternative* since the last food from the other alternative. This appears to be much more difficult. Not only is accuracy in sequential alternation tasks typically low (Kundey & Rowan, 2009), but patch-leaving in foraging analogues where each patch contains a fixed number of prey is generally a function of the time since the last prey capture and not a function of the number of consumed prey this patch (Krebs et al., 1974; Lima, 1984; Roche et al., 1998).

Further adding to the difficulty of the present task is that discrimination of the location of the previous reinforcer(s) required discrimination of the response(s) that immediately preceded those reinforcer(s). Although pigeons have proven capable of discriminating the most recent food-productive response (Alsop & Davison, 1992; Jones & Davison, 1998; Killeen, 1978), response discrimination does appear to be at least somewhat degraded by biologically relevant stimuli such as food (Killeen & Smith, 1984). Discriminating the responses emitted prior to a sequence of foods may be too difficult a task. Accurate performance in such a procedure may at least require some pre-training or a correction procedure not employed here.

The failure to discriminate the end of a sequence suggests that numerical discriminations based on the response preceding a reinforcer 2 to 5 reinforcers in the past may be beyond the abilities of a pigeon. Discriminating the start of a sequence may have been easier: In all conditions where reinforcers were arranged in sequences, a discontinuation signalled that the local probability of a further reinforcer on that alternative was 1.0. This was the case whether sequences were 2, 3 or 5 long. Despite this, preference was further towards the just-reinforced alternative when sequences were longer. This is perhaps the strongest evidence that there was no control by the sequences per se, and that the overall probability of a continuation controlled behaviour.

It appears that average patch size, though not its variance, directly controlled behaviour. In some sense, then, these pigeons displayed some numerical competence not identified by Davis and Perusse (1988). Those authors developed a classification system for numerical competence consisting of relative numerousness judgements ("Which array contains a greater number?"), subitizing (quick accurate estimates of small amounts), estimation (quick, less accurate estimates of amounts larger than 6), and what they identified as the most sophisticated skill: counting. The pigeons in Experiment 4 clearly did none of these. There was, however, an effect of the average number of same-alternative reinforcers. This suggests that numerical competence in nonhuman animals is perhaps more difficult to define. It appears that, as Johnson (1988) noted in commentary to Davis and Perusse's paper, nonhuman numerical competence can be quite different from human numerical competence. In the case of Experiment 4, it appears that there can be control by the average number of foods from an alternative in the absence of any counting, subtitizing, or estimating. These results support the view that relatively long-term aggregations of reinforcers can directly control behaviour. Although control by short-term contingencies is possible (as demonstrated in Experiments 2 and 3), it is not necessary for control by relatively longer-term aggregations of reinforcers.

Chapter VII

7.1 Summary

Throughout the experiments reported in this thesis, food and non-food stimuli (keylight colour changes) were delivered to pigeons contingent on keypeck responses. The effects of these response-contingent events on subsequent choice responding can be classified as resulting from the discriminative and the instrumental properties of the response-contingent events.

7.1.1 Discriminative functions of response-contingent events

When a response-contingent keylight illumination (Experiment 2) or food (Experiment 3) signalled a local left: right food ratio which differed from the overall food ratio, preference in the post-event period was clearly towards the locally richer alternative. This suggests a signalling or discriminative function of the response-contingent event: the event signalled what response (left vs. right keypeck) was more likely to produce the next food and local preference was shifted towards that locally richer alternative. Local preference was not, as the law of effect would predict, simply towards the just-productive alternative.

When the local food ratio signalled by a response-contingent event was equal to the overall food ratio, local preference in Experiments 1 and 2 was towards the just-productive alternative. At first glance, this may appear to suggest some direct strengthening function of the response-contingent event: although the next food was equally likely to come from the just-productive or the not-just-productive alternative, preference was reliably towards the just-productive alternative. Upon deeper analysis, the transient period of preference to the just-productive alternative was due to the post-event changeover requirement. The changeover delay in Experiment 1 made reinforcers on the not-just-responded-to (and therefore not-just-reinforced) alternative unavailable for the period of its duration. This temporary period of extinction was discriminable and worked to bias responding away from switching throughout the session, including the period immediately after a reinforced response. In Experiment 2 the local response-contingent event ratios throughout the post-keylight period were plotted. Although the local *arranged* food ratio may have been equal to the overall food ratio, the local *obtained* food ratio immediately after a response-contingent event was biased towards the just-productive alternative. This extreme local obtained food ratio then continued to maintain the extreme local behaviour ratio in a dynamical system: the extreme local behaviour ratio ensured a continuing extreme local reinforcer ratio which then continued to maintain the extreme behaviour ratio.

The bias-producing and enhancing effects of the post-event changeover contingencies were confirmed by directly manipulating them in Experiment 3. Local preference immediately after a response-contingent food was always clearly towards the alternative that did not require a changeover response to access, whether this alternative was the just-reinforced or the not-just-reinforced alternative, and despite the fact that the arranged food ratio in that period was 1:1.

When the reinforcers strictly alternated in Experiment 3, there was always some tendency to respond to the not-just-productive alternative. This adherence to the strict-alternation contingencies confirmed that differential control by the location of the response prior to the last reinforcer (nominally, the location of the last reinforcer) was possible. Thus local post-food preference in the random-alternation conditions of Experiment 3 was towards the alternative that was experienced as locally richer because of the changeover contingences. This may also be the case whenever the local arranged reinforcer ratio is 1:1 (including in Experiments 1 & 2): local preference after a response-contingent event is a function of the experienced, perceived local reinforcer ratio. Due to the influence of the changeover contingencies, these effective reinforcer ratios are not necessarily equivalent to the reinforcer ratios arranged by the experimenter. Thus, the transient period of extreme preference to the just-productive alternative, which initially appeared to indicate a strengthening function of the last response-contingent event, is, upon deeper analysis, actually indicative of a discriminative function. Response-contingent food and non-food stimuli guide preference towards the alternative that, in past similar situations, the animal has experienced as locally richer.

In Experiment 4 reinforcers appeared in fixed sequences of strictly (or randomly) alternating reinforcers. Preference was a function of the global probability of a same-alternative reinforcer (as in Krägeloh et al.'s, 2005 experiment) and there was no effect of the local probability of a same-alternative reinforcer (either 0 or 1). Thus, while a VI:EXT or EXT:VI schedule can be signalled by a single reinforcer (as in the strict alternation conditions of Experiment 3), it apparently cannot be signalled by a sequence of two or more reinforcers (as in Experiment 4). This lack of discriminative control by temporally distant reinforcers does not imply that such distant reinforcers have no effect on current behaviour. Clear control by temporally distant reinforcers has been previously reported a number of times (Aparicio & Baum, 2006; Baum & Davison, 2004; Landon et al., 2002; Experiment 3). These findings are not necessarily inconsistent with one another. A number of models consider current behaviour a dual function of the immediately preceding consequence and *behaviour* prior to that consequence (Baum & Davison, 2009; Bush & Mosteller, 1951; Davison & Hunter, 1979; Kacelnik et al., 1987; Killeen, 1984; Lea & Dow, 1984). As the behaviour prior to the most recent consequence was itself determined by earlier consequences, the term for earlier behaviour encapsulates the effects of these earlier consequences. If behaviour prior to the most recent reinforcer carries over (as found by Davison, Marr & Elliffe, in press), there need not be any discrimination of the location of earlier reinforcers in order for those reinforcers to have enduring effects.

The discriminative control that a response-contingent food or non-food event can acquire is subject to the perceptual or cognitive abilities of the target organism. If an organism simply cannot hold the location of successive response-contingent events in working memory, these multiple events will not acquire discriminative control (Experiment 4). Similarly, if the location of the last response-contingent keylight illumination is not retained at the time of the subsequent food delivery, there will be no discriminative control by that keylight illumination at that temporal distance (Experiment 2a).

7.1.2 Strengthening functions of response-contingent events

While local preference immediately after a food reinforcer pointed to food's discriminative functions (the alternative experienced as locally richer was preferred), preference in the inter-food interval pointed to food's strengthening functions. In Experiment 3 preference shifted further and further towards the just-productive alternative as the number of preceding same-alternative reinforcers increased. This increasing preference for the just-productive alternative was present regardless of the post-food changeover contingencies, and there was no similar effect of successive same-alternative post-food keylight illuminations. Additionally, there was no indication of increased preference to the not-just-productive alternative as the number of preceding strictly alternating (left, right, left, right, etc) reinforcers increased. Thus, the increased preference to the just-productive alternative as a function of increasing preceding continuations was likely due to some response-strengthening or instrumental function of the reinforcers. These response-strengthening properties were apparently masked by the rather large discriminative effects when preference was considered only as a function of the most recent reinforcer. They did exist however, and moreover were cumulative and enduring, becoming plainly visible in the lag analyses as well as the preference trees. The discriminative effects in contrast while larger in the short term, were short-lived.

7.2 Levels of analyses

The distinct discriminative and strengthening functions of response-contingent foods in Experiment 3 were only discovered because choice was analysed at different levels of temporal extendedness: the discriminative functions were visible at the level of behaviour throughout the post-food period (preference pulses) while the strengthening functions were visible at the level of the aggregated behaviour ratio across the entire post-food period as a function of the sequence of previous reinforcers (preference trees and lag analyses). Similarly, while pairing the red keylight stimuli with food in Experiment 1 had an effect on the post-stimulus preference pulse, the effect on the preference trees was small and inconsistent, and there was no effect whatsoever at the most extended level of analysis (generalized matching). Thus, analyses ought to be conducted at more than one level of temporal acuity to ensure that all response-contingent event effects are detected. Even a behaviour as seemingly simple as a pigeon pecking two keys for food reinforcement is apparently exceedingly complex and multiply controlled.

Williams' (1991a) found that when the local contingencies were clearly signalled, preference was towards the alternative momentarily more likely to provide the next reinforcer, while global measures of sensitivity to reinforcement were low. When the local contingencies were not clearly signalled, adherence to these local contingencies decreased and sensitivity to the global behaviour ratio increased. Thus, there can be control by contingencies operating at multiple levels of temporal extendedness. Moreover, control by the long-term distribution of reinforcers is not necessarily dependent on control at a local level (as is proposed by momentary maximization accounts). In the present experiments, when the local reinforcer ratio was clearly signalled, the local response ratio closely approximated the local reinforcer ratio (Experiment 2b & Experiment 3). When control by the local reinforcer ratio was degraded by arranging either a too-long stimulus-reinforcer delay (Experiment 2a) or a too-complicated discriminative stimulus (Experiment 4), preference was controlled by the long-term contingencies of reinforcement. In Experiment 2a, local preference after a response-contingent event did not fall to indifference, but to a level representative of the global ratio of the subsequent event (Landon et al., 2003a). In Experiment 4, local preference was a function of the global probability of a continuation reinforcer (Krägeloh et al., 2005; Lima, 1984). Thus, both the long-term and the more immediate consequences of behaviour were discriminated and the degree to which behaviour was controlled by contingencies at these two levels of temporal extendedness was a function of the degree to which the contingencies were discriminated.

Similarly Shahan and Podlesnik (2006; 2007) and Davison & Elliffe (2010) both reported a tradeoff in measured discrimination of elements in a two-element stimulus compound. When correct line orientation discriminations were reinforced at a higher rate than were correct colour discriminations, line-orientation discrimination exceeded colour discrimination (Shahan & Podlesnik, 2006), implying greater stimulus control by the stimulus element which better predicted reinforcement. In a similar way, control by extended and local contingencies may also be a function of not only the relative discriminabilities of these contingencies, but also the degree to which they predict reinforcement. The experiments in this thesis were not well-suited to addressing these questions, as control by the global contingencies could only be detected given weak control by the local contingencies. However, if control by contingencies at different levels of temporal extendedness is similar to control by contingencies signalled by different elements of a stimulus compound, then Shahan and Podlesnik's and Davison and Elliffe's research might imply that decreased control by contingencies at one level of temporal extendedness will be accompanied by increased control by contingencies at another level.

7.3 Conditional reinforcers

What of the non-food stimuli in Experiments 1 and 2? The results of Experiment 1 initially appeared to support a value-transfer account of conditional reinforcement. When examined in greater detail however, support for such an account weakened. Experiments 2 and 3 revealed that the period of increased preference to the just-productive alternative is better attributed to the changeover requirement making the local *obtained* food ratio VI:EXT or EXT:VI, thus biasing preference to the just-productive alternative. A question remains however as to why local post-stimulus preference in Experiment 1 became more extreme after the stimuli were paired with food.

When food followed a red keylight illumination in the paired conditions of Experiment 1, no additional response was required after the stimulus and before the food. However, this absence of a response requirement (after or during the stimulus presentation) may not have been discriminated. Prior to the paired conditions of Experiment 1, the pigeon subjects had extensive experience of food exclusively following pecks to a white keylight. Experiment 2 demonstrated (not for the first time; e.g., Wanchisen, 1990) that contingencies of reinforcement can have enduring effects which persist long after the contingencies themselves have been removed. Thus, the pigeons in Phase 2 of Experiment 1 may not have perceived the absence of a response requirement for the post-stimulus foods.

Experiment 2 confirmed that signalling an increase in the local reinforcer rate, combined with changeover contingencies in favour of staying and a conducive reinforcement history, can produce apparent conditional reinforcer effects. When examined in more detail (by plotting the local obtained reinforcer ratios and by arranging stimuli with different histories) these effects were attributed to a discriminative function of the stimuli. The effects of forward-pairing the stimuli with food in Experiment 1 may also have been due to some (perceived) change in the discriminative function of the stimuli, rather than to any change in their hedonic value. The results of Experiment 3 provide further evidence against a conditioned value interpretation of the Experiment 1 pairing effects. In Experiment 3, the post-food preference pulse was found to reflect the discriminative functions of the immediately preceding response-contingent event, while its strengthening or instrumental functions were better captured in the preference trees and lag analyses. The absence of any effect of pairing on the Experiment 1 post-stimulus preference trees, despite a clear effect on the preference pulse, suggests that pairing had some effect on the discriminative functions of the response-contingent stimuli, while the instrumental functions were unaffected.

Is there no place, then, for the strengthening, reinforcing effects of stimuli nominally termed conditional (or more commonly conditioned) reinforcers? The process of pairing an arbitrary stimulus with a stimulus that elicits a response is not a simple one. The effects are complicated and are, like the effects of unconditional reinforcers, not likely best summarized as either exclusively "reinforcing" or "discriminative". For example, Domjan and colleagues (reviewed in Domjan, Cusato, & Krause, 2004), working within a sexual conditioning paradigm, have found that the CR is more similar to the UR when the CS is similar to the US (Cusato & Domjan, 1998), that blocking and extinction are both attenuated when a more "naturalistic" CS is used (Koksal, Domjan, & Weisman, 1994; Krause, Cusato, & Domjan, 2003), and that these more naturalistic CSs better maintain responding when paired with neutral stimuli in second-order schedule procedures (Crawford & Domjan, 1995). Appropriate controls always confirm that both the arbitrary CS (a terrycloth stand which the male Japanese quail can grab, mount and make cloacal contact responses to) and the more naturalistic CS (the same terrycloth stand with a taxidermically prepared female quail head on top) are in fact initially neutral, in that no response is elicited by an unpaired stimulus. These results extend Garcia and Koelling's (1966) finding that rats learn a flavour-illness association with fewer pairings than are required to learn a light+sound-illness association, and will conversely more readily learn a light+sound-footshock pairing than a flavour-footshock pairing. Some stimulus-stimulus associations are prepared and the work of Domjan and colleagues extends this finding beyond the preparedness of one sensory modality over another.

Thus, although the stimulus must be paired with the US at some point in the individual's ontogenetic history in order for the response to follow a stimulus presentation, there is also a clear effect of phylogenetic history. Stimuli that have been historically paired with access to a receptive sexual partner have an enhanced ability (subsequent to pairing) to elicit behaviours appropriate to a sexual partner. This suggests an extended view of conditional reinforcement: not only is there an effect of what the stimulus has signalled about forthcoming reinforcement in the lifetime of the individual organism, but there is also clearly an effect of what that stimulus has signalled in the evolutionary history of that species.

7.4 The biological basis of learning

Davison and Baum (2006; 2010) and Shahan (2010) recently suggested that reinforcer effects may be regarded as primarily, or even entirely, discriminative. Although compelling, this strong version of a discriminative or signpost hypothesis is not generally supported by this thesis. The argument that reinforcers function by guiding behaviour towards the sources of further reinforcement would seem well-supported from a phylogenetic perspective: Any ability to discern the relationship between successive response-contingent events and to adjust one's behaviour accordingly would confer an obvious reproductive advantage, and would thus persist into successive generations. A tendency simply to repeat the just-reinforced response would likely be maladaptive, especially if in discordance with the current environment.

The above assumes an environment where the relationship between successive response-contingent events is ever-changing and thus unpredictable across generations, although somewhat predictable within a generation (Stephens, 1991).

However, the clear “bias” that the pigeons in Experiment 3 demonstrated towards repeating the previously-reinforced response suggests some inter-generational reliability in the relevant features of the environment. These animals may be in some sense “prepared” (Seligman, 1970) to learn that the just-productive response is likely to produce the next reinforcer. Thus, behaviour appropriate to such contingencies becomes apparent after only a few same-alternative continuations (which occurred just by chance). On the other hand, pigeons may be “unprepared” to learn that the not-just-productive response is more likely to produce the next reinforcer. Over the same period of a few discontinuations, there was no suggestion of behaviour approximating strict alternation. Saying that pigeons are unprepared to learn to avoid the just-productive response is not equivalent to saying that they cannot learn to behave in this way. Preference was clearly towards the not-just-productive alternative after a food in all of the strict alternation conditions of Experiment 3.

Thus, appetitive stimuli throughout pigeons’ evolutionary history were likely positively correlated with one another: repeating a just-successful response was more likely to result in further reinforcement than was avoiding such a response. While reinforcers may have generally been positively correlated with one another, this correlation may not have been very large and was likely characterized by high variability. Such an environment would select an ability to shift from a “repeat strategy” to an “avoid strategy” as a function of the current probability of a repeat reinforcer (versus the current probability of an avoid reinforcer), with some bias to adopting the “repeat” strategy or some enhanced ability to detect the contingencies appropriate to this strategy.

Cross-species comparisons may provide some empirical support for such historical-ecological hypotheses (Johnston, 1981). While pigeons apparently learn the behaviour appropriate to win-stay, lose-shift contingencies more readily than they learn the behaviour appropriate to win-shift, lose-stay contingencies (Randall & Zentall, 1997; Shimp, 1976b), the opposite appears to be the case for nectar-feeding hummingbirds (Cole, Hainsworth, Kamil, Mercier, & Wolf, 1982). This has been attributed to the fact that, while pigeons' prey tends to appear in "clumped" distributions (one prey predicts others in a similar location), once a flower has been visited, it is unlikely to provide further nectar. A test on an omnivorous species (noisy miners; Sulikowski & Burke, 2010) found that both the initial (Session 1 of Condition 1) and asymptotic accuracy on a win-shift task were always higher when nectar reinforcers were used than when invertebrate reinforcers were used. These results point to a historical-ecological influence on the tendency either to repeat or avoid a previously reinforced response: reinforcers typically arranged in clumped arrays support a tendency to repeat the just-reinforced response while reinforcers typically arranged singly support a tendency to avoid the just-reinforced response. Learning to emit the opposite response is not impossible. In all cases where the contrary response was required, the animals performed well above chance. Thus, the tendency to stay or to shift is by no means ingrained or immune from modification by consequences within the lifetime of an individual organism. Rather, the historical, species-typical environment seems to create some predisposition or preparedness, making some contingencies easier to learn than others.

Such references to historical, species-typical, inter-generational contingencies, while compelling, are necessarily post-hoc, and as such, may not be, in themselves, satisfying explanations for the tendency to repeat or avoid a successful response. Recent investigations have empirically demonstrated that learning is indeed related to regularities in the long-term, inter-generational environment (reviewed in Kawecki, 2010). In these studies, each generation of fruit fly is first exposed to an experience phase, where two oviposition substrates (orange and pineapple juice) are available. One of these oviposition substrates contains quinine (a substance which, although odourless, has a flavour unpleasant to fruit flies). All eggs laid in this phase are discarded. In the second phase (the consequence phase), the two oviposition substrates (neither containing quinine) are again presented. Subsequent to this phase, the eggs laid on the oviposition substrate formerly tainted with quinine are discarded, while the eggs laid on the previously untainted substrate are retained and form the next generation. The initial generation of fruit flies exposed to this procedure did not differ from an unselected control group in terms of the proportion of eggs laid in the never-tainted substrate (and were equally likely to lay their eggs in the previously tainted and never tainted substrate). However, within 30 generations, noticeable differences emerged, with the selected group becoming much more likely to lay their eggs in the never-tainted substrate (Dunlap & Stephens, 2009; Mery & Kawecki, 2002). This demonstrates that inter-generational regularities can shape the behavioural tendencies of individual organisms. Moreover, Mery and Kawecki also reported differences in the number of trials required for individual fruit flies from each group (selected or unselected) to learn to avoid a substrate paired with quinine (although learning was present in both groups). Additionally, the selected group avoided the tainted substrate after a delay of 3 hours, while the unselected control group fell to chance levels after 1 hour. Perhaps even more impressive, the selected fruit flies performed better than did the unselected controls on a different task: the fruit flies were put into a tube and exposed to two odours. Violent shocks accompanied (with a 4-s delay) one of the odours, while the other odour signalled shock-safety. After 15-20 minutes of exposure to these odour-shock relations, the fruit flies were put into a T-maze where they could choose between the two odours. The selected fruit flies were much more likely to choose the side of the T-maze which contained the odour never paired with shock (Mery, Pont, Preat, & Kawecki, 2007). These results, along with the cross-species comparisons of win-shift and win-stay behaviour, confirm that behaviour is a function of events in the historical, species-typical environment as well as events in the individual organism's environment.

7.5 Concluding comments

The 4 experiments of this thesis generalize and extend a number of findings. Experiments 1 and 2 clarified the situations in which response-contingent non-food stimuli can come to mimic the local effects of response-contingent food. Experiment 3 confirmed that pigeons are capable of behaving appropriately in strict alternation procedures and offered an explanation for the often-obtained better performance in win-stay preparations relative to win-shift preparations. Experiment 3 also reported an invariant tendency for preference to shift increasingly towards the just-reinforced alternative, suggesting some instrumental or strengthening function of reinforcers (which, as argued above, was likely due to an evolutionary history where repeat responses were often reinforced). Experiment 4 confirmed Krägeloh et al.'s (2005) finding that the *average* probability of a continuation controls local post-reinforcer preference and found that the local probability of a continuation had no effect on local preference.

Throughout this thesis, preference was a function of both the global and the local contingencies of reinforcement as signalled by response-contingent food and non-food events. This suggests some generalized ability to use environmental consistencies (in a variety of forms) to predict important biological events. Such an ability is clearly advantageous in an even somewhat inconsistent environment in which the predictors of important events can vary both within and across lifetimes. At core, the capacity to learn stimulus-response or stimulus-stimulus associations *must* exist because it provides some adaptive utility to the organism (Domjan, 2005; Domjan et al., 2004; Rozin & Kalat, 1971; Timberlake, 1993). Japanese quail with a history of CS presentations prior to access to a sexually receptive partner (the US) have greater reproductive success than do control animals with no prior such history (Mahometa & Domjan, 2005; Matthews, Domjan, Ramsey, & Crews, 2007). A similar result has also been obtained with male blue gourami fish: more offspring result when the mating opportunity is reliably preceded by an arbitrary stimulus (Hollis, Pharr, Dumas, Britton, & Field, 1997). Learning the environmental predictors for biologically relevant events has direct and clear reproductive advantages and thus *must* persist into subsequent generations. Pigeons exist in a variety of, often human-created, environments where the predictors of appetitive and aversive events are highly variable within and across generations, as well as across geographical space. That these animals thrive in a variety of environments no doubt speaks to their ability to discriminate a variety of relationships between events. Any tendency to simply repeat a previously successful response, regardless of the signalled consequences for doing so, is plainly maladaptive. Humans, too, exist and thrive in a variety of environments. While this has been attributed to our species' (unique) capacity for cultural innovation (Sterelny, 2006), such worldwide success at the species-level would no doubt be impossible without our (general) ability not only to discriminate inter-event relationships, but also to quickly adjust our behaviour accordingly (Krägeloh, Zapanta, Shepherd, & Landon, 2010; Lie, Harper, & Hunter, 2009).

Current behaviour is a function of prior experience with regularities between responses and their consequences. All of these influences can be parsimoniously understood within a framework that considers behaviour a function of contingencies operating at different levels of temporal extendedness. The interaction between the local behaviour and obtained food ratios in Experiment 2 demonstrated that preference was a function of the obtained food ratio in a particular two-second time bin. At a slightly more extended level, preference in the interval between successive response-contingent events in Experiments 2 and 3 was a function of the local food ratio in that period. At a more extended level still, an effect of multiple successive response-contingent foods was demonstrated in Experiments 3 and 4: preference was further towards the just-productive alternative given a longer preceding (Experiment 3) or forthcoming (Experiment 4) sequence of same-alternative foods. At an even more extended level, preference in Experiment 1 was clearly a function of each session's global left: right reinforcer ratio. Arguably, the enduring effects of historical contingencies seen in Experiment 2 demonstrate an even more extended level: the organism's lifetime may be considered the ultimate level of temporal extendedness. There is one further level however: contingencies operating at the level of the species' lifetime were evident in the apparent bias to repeat (rather than avoid) the just-reinforced response. Whether a response will be repeated or avoided is not a simple function of whether that response was followed by a reinforcer (the law of effect). Rather, the similarity of the current situation to earlier situations, and the previously experienced consequences for responding in these situations, determines whether a response will be repeated or avoided. Moreover, these earlier regularities exist at multiple levels of temporal acuity, ranging from seconds to months and even across evolutionary time.

References

Ahearn, W.H. (2003). Using simultaneous presentation to increase vegetable consumption in a mildly selective child with autism. *Journal of Applied Behavior Analysis, 36*, 361-365.

Allen, K. D., & Lattal, K. A. (1989). On conditioned reinforcing effects of negative discriminative stimuli. *Journal of the Experimental Analysis of Behavior, 52*, 335-339.

Alsop, B., & Davison, M., C. (1991). Effects of varying stimulus disparity and the reinforcer ratio in concurrent-schedule and signal-detection procedures. *Journal of the Experimental Analysis of Behavior, 56*, 67-80.

Alsop, B., & Davison, M. (1992). Discriminability between alternatives in a switching-key concurrent schedule. *Journal of the Experimental Analysis of Behavior, 57*, 51-65.

Aparicio, C. F., & Baum, W. M. (1997). Comparing locomotion with lever-press travel in an operant simulation of foraging. *Journal of the Experimental Analysis of Behavior, 68*, 177-192.

Aparicio, C. F., & Baum, W. M. (2006). Fix and sample with rats in the dynamics of choice. *Journal of the Experimental Analysis of Behavior, 86*(1), 43-63.

Aparicio, C. F., & Baum, W. M. (2009). Dynamics of choice: relative rate and amount affect local preference at three different time scales. *Journal of the Experimental Analysis of Behavior, 91*, 293-317.

Ayllon, T., & Azrin, N. H. (1966). Punishment as a discriminative stimulus and conditioned reinforcer with humans. *Journal of the Experimental Analysis of Behavior, 9*, 411-419.

Azrin, N. H., & Holz, W. C. (1966). Punishment. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application* (pp. 380-447). New York: Appleton-Century-Crofts.

Baily, J. T., & Mazur, J. E. (1990). Choice behavior in transition: development of preference for the higher probability of reinforcement. *Journal of the Experimental Analysis of Behavior, 53*, 409-422.

Balsam, P. D., Fairhurst, S., & Gallistel, C. R. (2006). Pavlovian contingencies and temporal information. *Journal of Experimental Psychology: Animal Behavior Processes, 32*, 284-294.

Baum, W. M. (1974). On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior, 22*(231-242).

Baum, W. M. (1979). Matching, undermatching, and overmatching in studies of choice. *Journal of the Experimental Analysis of Behavior, 32*, 269-281.

Baum, W. M. (1982). Choice, changeover, and travel. *Journal of the Experimental Analysis of Behavior, 38*, 35-49.

Baum, W. M., & Davison, M. (2004). Choice in a variable environment: Visit patterns in the dynamics of choice. *Journal of the Experimental Analysis of Behavior, 81*(1), 85-127.

Baum, W. M., & Davison, M. (2009). Modeling the dynamics of choice. *Behavioural Processes, 81*, 189-194.

Blanchard, R. (1975). The effect of S- on observing behavior. *Learning and Motivation, 6*, 1-10.

Brown, P. L., & Jenkins, H. M. (1968). Auto-shaping of the pigeon's key-peck. *Journal of the Experimental Analysis of Behavior, 11*, 1-8.

Buckner, R. L., Green, L., & Myerson, J. (1993). Short-term and long-term effects of reinforcers on choice. *Journal of the Experimental Analysis of Behavior, 59*, 293-307.

Bush, R. R., & Mosteller, F. (1951). A mathematical model for simple learning. *The Psychological review, 68*, 313-323.

Capaldi, E. J., & Miller, D. j. (1988a). Counting in rats: its functional significance and the independent cognitive processes that constitute it. *Journal of Experimental Psychology: Animal Behavior Processes, 14*, 3-17.

Capaldi, E. J., & Miller, D. j. (1988b). Number tags applied by rats to reinforcers are general and exert powerful control over responding. *The Quarterly Journal of Experimental Psychology, 40B*, 279-297.

Case, D. A., Ploog, B. O., & Fantino, E. (1990). Observing in a computer game. *Journal of the Experimental Analysis of Behavior, 54*, 185-199.

Cole, S., Hainsworth, F. R., Kamil, A. C., Mercier, T., & Wolf, L. L. (1982). Spatial-learning as an adaptation in hummingbirds. *Science, 217*, 655-657.

Cowie, R. J. (1977). Optimal foraging in Great Tits. *Nature, 268*, 137-139.

Crawford, L. L., & Domjan, M. (1995). Second-order sexual conditioning in male Japanese quail (*Coturnix japonica*). *Animal Learning & Behavior, 23*, 327-334.

Critchfield, T. S., & Lattal, K. A. (1993). Acquisition of a spatially defined operant with delayed reinforcement. *Journal of the Experimental Analysis of Behavior, 59*, 373-387.

Crombag, H. S., Galarce, E. M., & Holland, P. C. (2008). Pavlovian influences on goal-directed behavior in mice: the role of cue-reinforcer relations. *Learning & Memory, 15*, 299-303.

Cusato, B., & Domjan, M. (1998). Special efficacy of sexual conditioned stimuli that include species typical cues: tests with a conditioned stimuli preexposure design. *Learning and Motivation, 29*, 152-167.

Davis, H., & McIntire, R. W. (1969). Conditioned suppression under positive, negative, and contingency between conditioned and unconditioned stimuli. *Journal of the Experimental Analysis of Behavior, 12*, 633-640.

Davis, H., & Memmott, J. (1983). Autocontingencies: rats count to three to predict safety from shock. *Animal Learning & Behavior, 11*, 95-100.

Davis, H., & Perusse, R. (1988). Numerical competence in animals: definitional issues, current evidence, and a new research agenda. *Behavioral and Brain Sciences, 11*, 561-615.

Davison, M. (1991). Choice, changeover, and travel: a quantitative model. *Journal of the Experimental Analysis of Behavior, 55*, 47-61.

Davison, M. (1998). Experimental design: problems in understanding the dynamical behavior-environment system. *The Behavior Analyst, 21*, 219-240.

Davison, M., & Baum, W. M. (2000). Choice in a variable environment: Every reinforcer counts. *Journal of the Experimental Analysis of Behavior, 74*(1), 1-24.

Davison, M., & Baum, W. M. (2002). Choice in a variable environment: Effects of blackout duration and extinction between components. *Journal of the Experimental Analysis of Behavior, 77*(1), 65-89.

Davison, M., & Baum, W. M. (2003). Every reinforcer counts: Reinforcer magnitude and local preference. *Journal of the Experimental Analysis of Behavior, 80*(1), 95-129.

Davison, M., & Baum, W. M. (2006). Do conditional reinforcers count? *Journal of the Experimental Analysis of Behavior, 86*(3), 269-283.

Davison, M., & Baum, W. M. (2010). Stimulus effects on local preference: stimulus-response contingencies, stimulus-food pairing, and stimulus food correlation. *Journal of the Experimental Analysis of Behavior, 93*, 45-59.

Davison, M., & Elliffe, D. (2000). Travel time and concurrent-schedule choice: retrospective versus prospective control. *Journal of the Experimental Analysis of Behavior, 73*, 65-77.

Davison, M., & Elliffe, D. (2010). Divided stimulus control: a replication and a quantitative model. *Journal of the Experimental Analysis of Behavior*, 94, 13-23.

Davison, M., & Jenkins, P. E. (1985). Stimulus discriminability, contingency discriminability, and schedule performance. *Animal Learning & Behavior*(77-84).

Davison, M., Marr, M. J., & Elliffe, D. (in press). The effects of a local negative feedback function between choice and relative reinforcer rate. *Journal of the Experimental Analysis of Behavior*.

Davison, M., & McCarthy, D. (1988). *The Matching Law: A Research Review*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Davison, M., & Nevin, J. A. (1999). Stimuli, reinforcers, and behavior: an integration. *Journal of the Experimental Analysis of Behavior, 71*, 439-482.

Davison, M. C., & Hunter, I. W. (1979). Concurrent schedules: undermatching and control by previous experimental conditions. *Journal of the Experimental Analysis of Behavior, 32*, 233-244.

Dehaene, S. (1997). *The number sense: how the mind create mathematics*. London: Penguin Books.

Devonport, L., Hill, T., Wilson, M., & Ogden, E. (1997). Tracking and averaging invariable environments: a transition rule. *Journal of Experimental Psychology: Animal Behavior Processes, 23*, 450-460.

Devonport, L. D., & Devonport, J. A. (1994). Time-dependent averaging of foraging information in least chipmunks and golden-mantled ground squirrels. *Animal Behaviour, 47*, 787-802.

Dinsmoor, J. A. (1985). The role of observing and attention in establishing stimulus control. *Journal of the Experimental Analysis of Behavior, 43*, 365-381.

Dinsmoor, J. A. (2004). The etymology of basic concepts in the experimental analysis of behavior. *Journal of the Experimental Analysis of Behavior, 82*, 311-316.

Dinsmoor, J. A., Browne, M. P., & Lawrence, C. E. (1972). A test of the negative discriminative stimulus as a reinforcer of observing. *Journal of the Experimental Analysis of Behavior, 18*, 79-85.

Domjan, M. (2005). Pavlovian conditioning: a functional perspective. *Annual Review of Psychology, 56*, 179-206.

Domjan, M., Cusato, B., & Krause, M. (2004). Learning with arbitrary versus ecological conditioned stimuli: evidence from sexual conditioning. *Psychonomic Bulletin & Review, 11*, 232-246.

Dunlap, A. S., & Stephens, D. W. (2009). Components of change in the evolution of learning and unlearned preference. *Proceedings of the Royal Society B-Biological Sciences, 276*, 3201-3208.

Dunn, R., Williams, B. A., & Royalty, P. (1987). Devaluation of stimuli contingent on choice: evidence for conditioned reinforcement. *Journal of the Experimental Analysis of Behavior, 48*, 117-131.

Egger, M. D., & Miller, N. E. (1962). Secondary reinforcement in rats as a function of information value and reliability of the stimulus. *Journal of Experimental Psychology, 64*, 97-104.

Egger, M. D., & Miller, N. E. (1963). When is a reward reinforcing? An experimental study of the information hypothesis. *Journal of Comparative and Physiological Psychology, 56*, 132-137.

Egremont, P., & Rothschild, M. (1979). The calculating cormorants. *Biological Journal of the Linnean Society, 12*, 181-186.

Ellins, S. R., & von Kluge, S. (1990). Auditory food conditioning: effects of spatial contiguity and taste quality. *The Quarterly Journal of Experimental Psychology, 42B*, 73-86.

Escobar, R., & Bruner, C. A. (2007). Response induction during the acquisition and maintenance of lever pressing with delayed reinforcement. *Journal of the Experimental Analysis of Behavior, 88*, 49-49.

Ferster, C. B., & Hammer, C. (1965). Variables determining the effects of delay in reinforcement. *Journal of the Experimental Analysis of Behavior, 8*, 243-254.

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of Reinforcement*. New York: Appleton-Century-Crofts.

Fetterman, J. G., Stubbs, D. A., & Dreyfus, L. R. (1986). Scaling of events spaced in time. *Behavioural Processes, 13*, 53-68.

Freeman, T. J., & Lattal, K. A. (1992). Stimulus control of behavioral history. *Journal of the Experimental Analysis of Behavior, 57*, 5-15.

Galarce, E. M., Crombag, H. S., & Holland, P. C. (2007). Reinforcer-specificity of appetative and consummatory behavior of rats after Pavlovian conditioning with food reinforcers. *Physiology & Behavior, 91*(95-105).

Gallistel, C. R. (1990). *The Organization of Learning*. Cambridge, Massachusetts: The MIT Press.

Gallistel, C. R., King, A. P., Gottlieb, D., Balci, F., Papachristos, E. B., Szalecki, M., et al. (2007). Is matching innate? *Journal of the Experimental Analysis of Behavior, 87*, 161-199.

Gallistel, C. R., Mark, T. A., King, A. P., & Latham, P. E. (2001). The rat approximates an ideal detector of changes in rates of reward: implications for the law of effect. *Journal of Experimental Psychology: Animal Behavior Processes, 27*, 354-372.

Garcia, J., & Koelling, R. A. (1966). Relation of cue to consequence in avoidance learning. *Psychonomic Science, 4*, 123-124.

Gaynor, S. T., & Shull, R. L. (2002). The generality of selective observing. *Journal of the Experimental Analysis of Behavior, 77*, 171-187.

Gellermann, L., W. (1931). The double alternation problem: II. The behavior of children and human adults in a double alternation temporal maze. *Pedagogical Seminary and Journal of Genetic Psychology, 39*, 197-226.

Gibb, J. A. (1966). Tit predation and the abundance of Ernarmonia Conicolana (Heyl.) on Weeting Heath, Norfolk, 1962-1963. *Journal of Animal Ecology, 35*, 43-53.

Gibbon, J., & Balsam, P. (1981). Spreading association in time. In C. M. Locurto, H. S. Terrace & J. Gibbon (Eds.), *Autoshaping and Conditioning Theory* (pp. 219-253). New York: Academic Press.

Godfrey, R., & Davison, M. (1998). Effects of varying sample- and choice- stimulus disparity on symbolic matching-to-sample performance. *Journal of the Experimental Analysis of Behavior, 69*, 311-326.

Gollub, L. R. (1977). Conditioned reinforcement: schedule effects. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior*. Englewood Cliffs, N J: Prentice-Hall.

Grace, R. C. (1994). A contextual model of concurrent chains choice. *Journal of the Experimental Analysis of Behavior, 61*, 113-129.

Grant, D. S. (1975). Proactive interference in pigeon short-term memory. *Journal of Experimental Psychology: Animal Behavior Processes, 104*, 207-220.

Grant, D. S., & Roberts, W. A. (1973). memory trace interaction in pigeon short-term memory. *Journal of Experimental Psychology, 101*, 21-29.

Hearst, E. (1962). Delayed alternation in the pigeon. *Journal of the Experimental Analysis of Behavior, 5*, 225-228.

Hendry, D. P. (1969). Reinforcing value of information: fixed-ratio schedules. In D. P. Hendry (Ed.), *Conditioned Reinforcement*. Homewood, Illinois: The Dorsey Press.

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior, 4*, 267-272.

Herrnstein, R. J. (1964). Secondary reinforcement and rate of primary reinforcement. *Journal of the Experimental Analysis of Behavior, 7*, 27-36.

Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior, 13*, 243-266.

Herrnstein, R. J., & Loveland, D. H. (1975). Maximizing and matching on concurrent ratio schedules. *Journal of the Experimental Analysis of Behavior, 24*, 107-116.

Hinson, J. M., & Staddon, J. E. R. (1983). Matching, maximizing, and hill-climbing. *Journal of the Experimental Analysis of Behavior, 40*, 321-331.

Hirai, S., & Jitsumori, M. (2009). Counting absolute number of items, from 1 to 8, in pigeons. *Learning & Behavior, 37*, 365-379.

Hiraoka, K. (1984). Discrete-trial probability learning in rats: effects of local contingencies of reinforcement. *Learning & Behavior, 12*, 343-349.

Hollis, K. L., Pharr, V. L., Dumas, M. J., Britton, G. B., & Field, J. (1997). Classical conditioning provides paternity advantage for territorial male blue gouramis (*Trichogaster trichopterus*). *Journal of Comparative Psychology, 111*, 219-225.

Holz, W. C., & Azrin, N. H. (1961). Discriminative properties of punishment. *Journal of the Experimental Analysis of Behavior, 4*, 225-232.

Honig, W. K. (1978). Studies of working memory in the pigeon. In S. H. Hulse, H. Fowler & W. K. Honig (Eds.), *Cognitive Processes in Animal Behavior* (pp. 211-248). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Hunter, I., & Davison, M. (1985). Determination of a behavioral transfer function: white-noise analysis of session-to-session response ratio dynamics on concurrent VI VI schedules. *Journal of the Experimental Analysis of Behavior, 43*, 43-59.

Jenkins, H. M., Barnes, R. A., & Barrera, F. J. (1981). Why autoshaping depends on trial spacing. In C. M. Locurto, H. S. Terrace & J. Gibbon (Eds.), *Autoshaping and Conditioning Theory* (pp. 255-284). New York: Academic Press.

Jenkins, H. M., & Boakes, R. A. (1973). Observing stimulus sources that signal food or no food. *Journal of the Experimental Analysis of Behavior, 20*, 197-207.

Johnson, L. M., Bickel, W. K., Higgins, S. T., & Morris, E. K. (1991). The effects of schedule history and the opportunity for adjunctive responding on behavior during a fixed-interval schedule of reinforcement. *Journal of the Experimental Analysis of Behavior, 55*, 313-322.

Johnson, M. (1988). Out for the count. *Behavioral and Brain Sciences, 11*, 561-615.

Johnston, T. D. (1981). Contrasting approaches to a theory of learning. *Behavioral and Brain Sciences, 4*, 125-138.

Jones, B. M., & Davison, M. (1997). The control of choice by its consequences. *Journal of the Experimental Analysis of Behavior, 68*, 329-348.

Jones, B. M., & Davison, M. (1998). Reporting contingencies of reinforcement in concurrent schedules. *Journal of the Experimental Analysis of Behavior, 69*, 161-183.

Kacelnik, A., Krebs, J. R., & Ens, B. (1987). Foraging in a changing environment: an experiment with starlings (Sturnis Vulgaris). In M. L. Commons, A. Kacelnik & S. J. Shettleworth (Eds.), *Quantitative Analyses of Behavior: Foraging* (Vol. 6). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Kamil, A. C., & Roitblat, H. L. (1985). The ecology of foraging behavior: implications for animal learning and memory. *Annual Review of Psychology, 36*, 141-169.

Kamin, L. J. (1969). Predictability, surprise, attention and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 279-296). New York: Appleton-Century-Crofts.

Katz, H. N. (1976). A test of the reinforcing properties of stimuli correlated with nonreinforcement. *Journal of the Experimental Analysis of Behavior, 26*, 45-65.

Kawecki, T. J. (2010). Evolutionary ecology of learning: insights from fruit flies. *Population Ecology, 52*, 15-25.

Keely, J., Feola, T., & Lattal, K. A. (2007). Contingency tracking during unsignaled delayed reinforcement. *Journal of the Experimental Analysis of Behavior, 88*, 229-247.

Kelleher, R. T. (1966). Conditioned reinforcement in second-order schedules. *Journal of the Experimental Analysis of Behavior, 9*, 475-485.

Kelleher, R. T., & Gollub, L. R. (1962). A review of positive conditioned reinforcement. *Journal of the Experimental Analysis of Behavior, 5*, 543-597.

Kern, L., & Marder, T. J. (1996). A comparison of simultaneous and delayed reinforcement as treatments for food selectivity. *Journal of Applied Behavior Analysis, 29*, 243-246.

Killeen, P.R. (1978). Superstition: a matter of bias, not detectability. *Science, 199*, 88-90.

Killeen, P. R. (1982). Incentive theory II: models for choice. *Journal of the Experimental Analysis of Behavior, 38*, 217-232.

Killeen, P. R. (1984). Incentive theory III: adaptive clocks. *Annals of the New York Academy of Sciences, 423*, 515-527.

Killeen, P. R., & Smith, J. P. (1984). Perception of contingency in conditioning: scalar timing, response bias, and erasure of memory by reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes, 10*, 333-345.

Koksal, F., Domjan, M., & Weisman, G. (1994). Blocking of the sexual conditioning of differentially effective conditioned stimulus objects. *Animal Learning & Behavior, 22*, 103-111.

Koops, M. A. (2004). Reliability and the value of information. *Animal Behaviour, 67*, 103-111.

Krägeloh, C. U., & Davison, M. (2003). Concurrent-schedule performance in transition: Changeover delays and signaled reinforcer ratios. *Journal of the Experimental Analysis of Behavior, 79*(1), 87-109.

Krägeloh, C. U., Elliffe, D., M., & Davison, M. (2006). Contingency discriminability and peak shift in concurrent schedules. *Journal of the Experimental Analysis of Behavior, 86*(1), 11-30.

Krägeloh, C. U., Davison, M., & Elliffe, D., M. (2005). Local preference in concurrent schedules: The effects of reinforcer sequences. *Journal of the Experimental Analysis of Behavior, 84*(1), 37-64.

Krägeloh, C. U., Zapanta, A. E., Shepherd, D., & Landon, J. (2010). Human choice behaviour in a frequently changing environment. *Behavioural Processes, 83*, 119-126.

Krause, M. A., Cusato, B., & Domjan, M. (2003). Extinction of conditioned sexual responses in male Japanese quail (*Coturnix japonica*): Role of species-typical cues. *Journal of Comparative Psychology, 117*, 76-86.

Krebs, J. R., Ryan, J. C., & Charnov, E. L. (1974). Hunting by expectation or optimal foraging? A study of patch use by chickadees. *Animal Behaviour, 22*, 953-964.

Kundey, S. M. A., & Rowan, J. D. (2009). Single and double alternation learning in rats: the role of set size and correction. *Learning and Motivation, 40*, 1-14.

Landon, J., & Davison, M. (2001). Reinforcer-ratio variation and its effects on rate of adaptation. *Journal of the Experimental Analysis of Behavior, 75*(2), 207-234.

Landon, J., Davison, M., & Elliffe, D. (2002). Concurrent schedules: Short- and long-term effects of reinforcers. *Journal of the Experimental Analysis of Behavior, 77*(3), 257-271.

Landon, J., Davison, M., & Elliffe, D. (2003a). Choice in a variable environment: Effects of unequal reinforcer distributions. *Journal of the Experimental Analysis of Behavior, 80*(2), 187-204.

Landon, J., Davison, M., & Elliffe, D. (2003b). Concurrent Schedules: Reinforcer magnitude effects. *Journal of the Experimental Analysis of Behavior, 79*(3), 351-365.

Lattal, K. A. (1979). Reinforcement contingencies as discriminative stimuli: II. Effects of changes in stimulus probability. *Journal of the Experimental Analysis of Behavior, 31*, 15-22.

Lattal, K. A., & Gleeson, S. (1990). Response acquisition with delayed reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes, 16*, 27-39.

Lau, B., & Glimcher, P., W. (2005). Dynamic response-by-response models of matching behavior in Rhesus monkeys. *Journal of the Experimental Analysis of Behavior, 84*(3), 555-579.

Lea, S. E. G., & Dow, S. M. (1984). The integration of reinforcements over time. *Annals of the New York Academy of Sciences, 423*, 269-277.

Lewis, P., Lewin, L., Muehleisen, P., & Stoyak, M. (1974). Preference for signalled reinforcement. *Journal of the Experimental Analysis of Behavior, 22*, 143-150.

Lie, C., Harper, D. N., & Hunt, M. (2009). Human performance on a two-alternative rapid-acquisition task. *Behavioural Processes, 81*, 244-249.

Lieberman, D. A., Cathro, J. S., Nichol, K., & Watson, E. (1997). The role of S- in human observing behavior: bad news is sometimes better than no news. *Learning and Motivation, 28*, 20-42.

Lima, S. L. (1984). Downy Woodpecker foraging behavior: efficient sampling in simple and stochastic environments. *Ecology, 65*, 166-174.

Lobb, B., & Davison, M. C. (1975). Performance in concurrent interval schedules: a systematic replication. *Journal of the Experimental Analysis of Behavior, 24*, 191-197.

Mackintosh, N. J. (1973). Stimulus-selection: learning to ignore stimuli that predict no change in reinforcement. In R. A. Hinde & J. Stevenson-Hinde (Eds.), *Constraints on Learning: Limitations and Predispositions* (pp. 75-100). London: Academic Press.

Mackintosh, N. J. (1983). *Conditioning and Associative Learning*. Oxford: Claredon Press.

Mahometa, M. J., & Domjan, M. (2005). Classical conditioning increases reproductive success in Japanese quail, *Coturnix japonica*. *Animal Behaviour, 69*, 983-989.

Mark, T. A., & Gallistel, C. R. (1994). Kinetics of matching. *Journal of Experimental Psychology: Animal Behavior Processes, 20*, 79-95.

Marr, M. J. (1979). Second-order schedules and the generation of unitary response sequences. In M. D. Zeiler & P. Harzam (Eds.), *Advances in analysis of behaviour* (Vol. 1: Reinforcement and the organization of behaviour). Chichester, England: Wiley.

Matthews, R. N., Domjan, M., Ramsey, M., & Crews, D. (2007). Learning effects on sperm competition and reproductive fitness. *Psychological Science, 18*, 758-762.

Mazur, J. E. (1992). Choice behavior in transition: development of preference with ratio and interval schedules. *Journal of Experimental Psychology: Animal Behavior Processes, 18*, 364-378.

Mazur, J. E. (1997). Effects of rate of reinforcement and rate of change on choice behaviour in transition. *The Quarterly Journal of Experimental Psychology, 50B*, 111-128.

Mazur, J. E. (2001). Hyperbolic value addition and general models of animal choice. *Psychological Review, 108*, 96-112.

McLinn, C. M., & Stephens, D. W. (2006). What makes information valuable: signal reliability and environmental uncertainty. *Animal Behaviour, 71*, 1119-1129.

McMillan, D. E., & Morse, W. H. (1967). Schedules using noxious stimuli. II. Low intensity electric shock as a discriminative stimulus. *Journal of the Experimental Analysis of Behavior, 10*, 109-118.

McNamara, J. M., Green, R. F., & Olsson, O. (2006). Bayes' theorem and its applications in animal behaviour. *OIKOS, 112*, 243-251.

Mery, F., & Kawecki, T. J. (2002). Experimental evolution of learning ability in fruit flies. *Proceedings of the National Academy of Sciences of the United States of America, 99*, 14274-14279.

Mery, F., Pont, J., Preat, T., & Kawecki, T. J. (2007). Experimental evolution of olfactory memory in *Drosophila melanogaster*. *Physiological and Biochemical Zoology, 80*, 399-405.

Miller, J. T., Saunders, S. S., & Bourland, G. (1980). The role of stimulus disparity in concurrently available reinforcement schedules. *Animal Learning & Behavior, 8*, 635-641.

Miller, R. R., Barnet, R. C., & Grahame, N. J. (1995). Assessment of the Rescorla-Wagner model. *Psychological Bulletin, 117*, 363-386.

Mueller, K. L., & Dinsmoor, J. A. (1984). Testing the reinforcing properties of S-: a replication of Lieberman's procedure. *Journal of the Experimental Analysis of Behavior, 41*, 17-25.

Neuman, P., Ahearn, W. H., & Hineline, P. N. (2000). Pigeons' choices between fixed-ratio and linear or geometric escalating schedules. *Journal of the Experimental Analysis of Behavior, 73*, 93-102.

Neuringer, A. J., & Chung, S.-H. (1967). Quasi-reinforcement: control of responding by a percentage reinforcement schedule. *Journal of the Experimental Analysis of Behavior, 10*, 45-54.

Nevin, J. A. (1969). Interval reinforcement of choice behavior in discrete trials. *Journal of the Experimental Analysis of Behavior, 12*, 875-885.

Nevin, J. A., & Grace, R. C. (2000). Behavioral momentum and the law of effect. *Behavioral and Brain Sciences, 23*, 73-130.

Odum, A. L., Ward, R. D., Barnes, C. A., & Burke, K. A. (2006). The effects of delayed reinforcement on variability and repetition of response sequences. *Journal of the Experimental Analysis of Behavior 86*, 159-179.

Olsson, O., & Brown, J. S. (2006). The foraging benefits of information and the penalty of ignorance. *OIKOS, 112*, 260-273.

Olsson, O., & Holmgren, N. M. A. (1998). The survival-rate-maximizing policy for Bayesian foragers: wait for good news. *Behavioral Ecology, 9*, 345-353.

Olton, D. S. (1978). Characteristics of spatial memory. In S. H. Hulse, H. Fowler & W. K. Honig (Eds.), *Cognitive Processes in Animal Behavior* (pp. 341-373). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Palya, W., L, & Allan, R., W. (2003). Dynamical concurrent schedules. *Journal of the Experimental Analysis of Behavior, 79*(1), 1-20.

Papini, M. R., & Bitterman, M. E. (1990). The role of contingency in classical conditioning. *Psychological Review, 97*, 396-403.

Pavlov, I. P. (1904). Physiology of Digestion (Nobel Lecture). Retrieved November 11, 2009, from http://nobelprize.org/nobel_prizes/medicine/laureates/1904/pavlov-lecture.html

Pavlov, I. P. (1927). *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*. Oxford: Oxford University Press.

Pearce, J. M., & Bouton, M. E. (2001). Theories of associative learning in animals. *Annual Review of Psychology, 52*, 111-139.

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: variations in the effectiveness of conditioned but not unconditioned stimuli. *Psychological Review, 87*, 532-552.

Perone, M., & Baron, A. (1980). Reinforcement of human observing behavior by a stimulus correlated with extinction of increased effort. *Journal of the Experimental Analysis of Behavior, 34*, 239-261.

Perone, M., & Kaminski, B. J. (1992). Conditioned reinforcement of human observing behavior by descriptive and arbitrary verbal stimuli. *Journal of the Experimental Analysis of Behavior, 58*, 557-575.

Piazza, C. C., Patel, M. A., Santana, C. M., Goh, H.-L., Delia, M. D., & Lancaster, B. M. (2002). An evaluation of simultaneous and sequential presentation of preferred and nonpreferred food to treat food selectivity. *Journal of Applied Behavior Analysis, 35*, 259-270.

Pliskoff, S. S., Cicerone, R., & Nelson, T. D. (1978). Local response-rate constancy on concurrent variable-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior, 29*, 431-446.

Podlesnik, C. A., Jimenez-Gomez, C., Ward, R. D., & Shahan, T. A. (2009). Resistance to change and frequency of response-dependent stimuli uncorrelated with reinforcement. *Journal of the Experimental Analysis of Behavior, 92*, 199-214.

Purdy, J. E., & Peel, J. L. (1988). Observing response in goldfish (*Carassius auratus*). *Journal of Comparative Psychology, 102*, 160-168.

Randall, C., K, & Zentall, T. R. (1997). Win-stay/lose-shirt and win-shift/lose-stay learning by pigeons in the absence of overt response mediation. *Behavioural Processes, 41*, 227-236.

Reed, P., & Hall, G. (1989). The quasi-reinforcement effect: The influence of brief stimuli uncorrelated with reinforcement on variable ratio schedules. *Learning and Motivation, 20*, 242-261.

Rescorla, R. A. (1966). Predictability and number of pairings in Pavlovian fear conditioning. *Psychonomic Science*(4), 383-384.

Rescorla, R. A. (1967). Pavlovian conditioning and its proper control procedures. *Psychological Review, 74*, 71-80.

Rescorla, R. A. (1972). Informational variables in Pavlovian conditioning. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol. 6). New York: Academic Press.

Rescorla, R. A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology, 66*, 1-5.

Rescorla, R. A. (1980). Simultaneous and successive associations in sensory preconditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 6*, 207-216.

Rescorla, R. A. (1988). Pavlovian conditioning: it's not what you think it is. *American Psychologist, 43*, 151-160.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical Conditioning II: Current research and theory*. New York: Appleton-Century-Crofts.

Richards, R. W. (1981). A comparison of signaled and unsignaled delay of reinforcement. *Journal of the Experimental Analysis of Behavior, 35*, 145-152.

Riordan, M., Iwata, B. A., Finney, J. W., Wohl, M. K., & Stanley, A. E. (1984). Behavioral assessment and treatment of chronic food refusal in handicapped children. *Journal of Applied Behavior Analysis, 17*, 327-341.

Roberts, W. A. (1991). Testing optimal foraging theory on the radial maze - the role of learning in patch sampling. *Animal Learning & Behavior, 19*, 305-316.

Roberts, W. A., & Grant, D. S. (1976). Studies of short-term memory in the pigeon using the delayed matching to sample procedure. In D. L. Medin, W. A. Roberts & R. T. Davis (Eds.), *Processes of Animal Memory* (pp. 79-112). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

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.

Roche, J. P., Timberlake, W., Glanz, W. E., & Stubbs, D. A. (1998). The influence of current-visit experience within a prey patch on patch persistence. *Behavioural Processes, 43*, 11-25.

Rozin, P., & Kalat, J. W. (1971). Specific hungers and poison avoidance as adaptive specializations of learning. *Psychological Review, 78*, 459-486.

Sargisson, R. J., & White, K. G. (2001). Generalization of delayed matching to sample following training at different delays. *Journal of the Experimental Analysis of Behavior, 75*, 1-14.

Savastano, H. I., Cole, R. P., Barnet, R. C., & Miller, R. R. (1999). Reconsidering conditioned inhibition. *Learning and Motivation, 30*, 101-127.

Schaal, D. W., & Branch, M. N. (1988). Responding of pigeons under variable-interval schedules of unsignaled, and completely signaled delays to reinforcement. *Journal of the Experimental Analysis of Behavior, 50*, 33-54.

Schneider, S. M., & Davison, M. (2005). Demarcated response sequences and generalised matching. *Behavioural Processes, 70*, 51-61.

Schofield, G., & Davison, M. (1997). Nonstable concurrent choice in pigeons. *Journal of the Experimental Analysis of Behavior, 68*, 219-232.

Schrier, A. M., Thompson, C. R., & Spector, N. R. (1980). Observing Behavior in Monkeys (*Macaca arctoides*): Support for the information hypothesis. *Learning and Motivation, 11*, 355-365.

Seligman, M. E. P. (1970). On the generality of the laws of learning. *Psychological Review, 77*, 406-418.

Seligman, M. E. P., & Meyer, B. (1970). Chronic fear and ulceration in rats as a function of unpredictability of safety. *Journal of Comparative and Physiological Psychology, 73*, 202-207.

Shahan, T. A. (2010). Conditioned reinforcement and response strength. *Journal of the Experimental Analysis of Behavior, 93*, 269-289.

Shahan, T. A., & Lattal, K. A. (1998). On the functions of the changeover delay. *Journal of the Experimental Analysis of Behavior, 69*, 141-160.

Shahan, T. A., Magee, A., & Dobberstein, A. (2003). The resistance to change of observing. *Journal of the Experimental Analysis of Behavior, 80*, 273-293.

Shahan, T. A., & Podlesnik, C. A. (2005). Rate of conditioned reinforcement affects observing rate but not resistance to change. *Journal of the Experimental Analysis of Behavior, 84*, 1-17.

Shahan, T. A., & Podlesnik, C. A. (2006). Divided attention performance and the matching law. *Learning & Behavior, 34*, 255-261.

Shahan, T. A., & Podlesnik, C. A. (2007). Divided attention and the matching law: sample duration affectst sensitivity to reinforcement allocation. *Learning & Behavior, 35*, 141-148.

Shahan, T. A., & Podlesnik, C. A. (2008a). Conditioned reinforcement value and resistance to change. *Journal of the Experimental Analysis of Behavior, 89*, 263-298.

Shahan, T. A., & Podlesnik, C. A. (2008b). Quantitative analyses of observing and attending. *Behavioural Processes, 78*, 145-157.

Shimp, C. P. (1966). Probabilistically reinforced choice behavior in pigeons. *Journal of the Experimental Analysis of Behavior, 9*, 443-455.

Shimp, C. P. (1976a). Short-term memory in the pigeon: relative recency. *Journal of the Experimental Analysis of Behavior, 25*, 55-61.

Shimp, C. P. (1976b). Short-term memory in the pigeon: the previously reinforced response. *Journal of the Experimental Analysis of Behavior, 26*, 487-493.

Shull, R. L., & Pliskoff, S. S. (1967). Changeover delay and concurrent schedules - some effects on relative performance measures. *Journal of the Experimental Analysis of Behavior, 10*, 517-527.

Sidman, M. (1960). *Tactics of Scientific Research*. Boston: Authors Cooperative, Inc., Publishers.

Silberberg, A., Hamilton, B., Ziriax, J. M., & Casey, J. (1978). The structure of choice. *Journal of Experimental Psychology: Animal Behavior Processes, 4*, 368-398.

Silberberg, A., & Williams, D. R. (1974). Choice behavior on discrete trials: a demonstration of the occurrence of a response strategy. *Journal of the Experimental Analysis of Behavior, 21*, 315-322.

Skinner, B. F. (1938). *The Behavior of Organisms: An Experimental Analysis*. New York :: Appleton-Century-Crofts.

Solberg, K. M., Hanley, G. P., Layer, S. A., & Ingvarsson, E. T. (2007). The effects of reinforcer pairing and fading on preschoolers' snack selections. *Journal of Applied Behavior Analysis, 40*, 633-644.

Squires, N., & Fantino, E. (1971). A model for choice in simple concurrent and concurrent-chains schedules. *Journal of the Experimental Analysis of Behavior, 15*, 27-38.

Stefurak, T. L., Martin, G., & van der Kooy, D. (1990). The representation in memory of morphine's unconditioned motivational effects depends on the nature of the conditioned stimulus. *Psychobiology, 18*, 435-442.

Stephens, D. W. (1991). Change, regularity, and value in the evolution of animal learning. *Behavioral Ecology, 2*, 77-89.

Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory*. Princeton, New Jersey: Princeton University Press.

Sterelny, K. (2006). The evolution and evolvability of culture. *Mind & Language, 21*, 137-165.

Stubbs, D. A. (1971). Second-order schedules and the problem of conditioned reinforcement. *Journal of the Experimental Analysis of Behavior, 16*, 289-313.

Stubbs, D. A., & Cohen, S. L. (1972). Second-order schedules: comparison of different procedures for scheduling paired and nonpaired brief stimuli. *Journal of the Experimental Analysis of Behavior, 18*, 403-413.

Stubbs, D. A., & Silverman, P. J. (1972). Second-order schedules: brief shock at the completion of each component. *Journal of the Experimental Analysis of Behavior, 17*, 201-212.

Sulikowski, D., & Burke, D. (2010). Reward type influences performance and search structure of an omnivorous bird in an open-field maze. *Behavioural Processes, 83*, 31-35.

Sullivan, L. G. (1984). Long-delay and selective association in food aversion learning. *Quarterly Journal of Experimental Psychology, 36B*, 65-87.

Temple, W., Scown, J. M., & Foster, T. M. (1995). Changeover delay and concurrent-schedule performance in domestic hens. *Journal of the Experimental Analysis of Behavior, 63*, 71-95.

Tiger, J. H., & Hanley, G. P. (2006). Using reinforcer pairing and fading to increase the milk consumption of a preschool child. *Journal of Applied Behavior Analysis, 39*, 399-403.

Timberlake, W. (1993). Behavior systems and reinforcement: an integrative approach. *Journal of the Experimental Analysis of Behavior, 60*, 105-128.

Valone, T. J. (1992). Information for patch assessment - a field investigation with black-chinned humingbirds. *Behavioral Ecology, 3*, 211-222.

Valone, T. J. (2006). Are animals capable of Bayesian updating? An empirical review. *OIKOS, 112*, 252-259.

Wald, B. A., & Dukich, T. D. (1978). Observing behavior: redundant stimuli and time since information. *Animal Learning & Behavior, 6*, 380-384.

Wagner, A. R., Logan, K. H., & Price, T. (1968). Stimulus selection in animal discrimination learning. *Journal of Experimental Psychology, 76*, 171-180.

Wanchisen, B. A. (1990). Forgetting the lessons of history. *The Behavior Analyst, 13*, 31-37.

Wanchisen, B. A., Tatham, T. A., & Hineline, P. N. (1988). Pigeons' choices in situations of diminishing returns: fixed- versus progressive-ratio schedules. *Journal of the Experimental Analysis of Behavior, 50*, 375-394.

Wasserman, E. A. (1974). Stimulus-reinforcer predictiveness and selective discrimination learning in pigeons. *Journal of Experimental Psychology, 103*, 284-297.

Wasserman, E. A. (1990). Attribution of causality to common and distinctive elements of compound stimuli. *Psychological Science, 1*, 298-302.

Wasserman, E. A., & Miller, R. R. (1997). What's elementary about associative learning? *Annual Review of Psychology, 48*, 573-607.

Weiner, H. (1964). Conditioning history and human fixed-interval performance. *Journal of the Experimental Analysis of Behavior, 7*, 383-385.

Weiner, H. (1969). Controlling human fixed-interval performance. *Journal of the Experimental Analysis of Behavior, 12*, 349-373.

Weiner, H. (1970). Human behavioral persistence. *The Psychological Record, 20*, 445-456.

White, K. G. (2001). Forgetting functions. *Animal Learning & Behavior, 29*, 193-207.

White, K. G., & Cooney, E. B. (1996). Consequences of remembering: independence of performance at different retention intervals. *Journal of Experimental Psychology: Animal Behavior Processes, 22*, 51-59.

Williams, A. M., & Lattal, K. A. (1999). The role of the response-reinforcer relation in delay-of-reinforcement effects. *Journal of the Experimental Analysis of Behavior, 71*, 187-194.

Williams, B. A. (1971a). Color alternation learning in the pigeon under fixed-ratio schedules of reinforcement. *Journal of the Experimental Analysis of Behavior, 15*, 129-140.

Williams, B. A. (1971b). Non-spatial delayed alternation by the pigeon. *Journal of the Experimental Analysis of Behavior, 16*, 15-21.

Williams, B. A. (1976). Double alternation learning in pigeons. *Journal of General Psychology, 94*, 285-293.

Williams, B. A. (1991a). Choice as a function of local versus molar reinforcement contingencies. *Journal of the Experimental Analysis of Behavior, 56*, 455-473.

Williams, B. A. (1991b). Marking and bridging versus conditioned reinforcement. *Animal Learning & Behavior, 19*, 264-269.

Williams, B. A. (1994). Conditioned reinforcement: neglected or outmoded explanatory construct? *Psychonomic Bulletin & Review, 1*, 457-475.

Williams, B. A. (2003). Challenges to timing-based theories of operant behavior. *Behavioural Processes, 62*, 115-123.

Williams, D. R., & Williams, H. (1969). Auto-maintenance in the pigeon: sustained pecking despite contingent non-reinforcement. *Journal of the Experimental Analysis of Behavior, 12*, 511-520.

Wilton, R. N., & Clements, R. O. (1971). Observing responses and informative stimuli. *Journal of the Experimental Analysis of Behavior, 15*, 199-204.

Woodard, W. T., Ballinger, J. C., & Bitterman, M. E. (1974). Autoshaping: further study of "negative automaintenance". *Journal of the Experimental Analysis of Behavior, 22*, 47-51.

Wright, A. A. (2007). An experimental analysis of memory processing. *Journal of the Experimental Analysis of Behavior, 88*, 405-433.

Zeiler, M. D. (1987). On optimal choice strategies. *Journal of Experimental Psychology: Animal Behavior Processes, 13*, 31-39.