Reinforcer devaluation affects within-bout responses to a greater degree than bout-initiation responses on human free-operant responding

Phil Reed[1]

[1]Swansea University, UK

Correspondence address: Phil Reed,
Department of Psychology,
Swansea University.
Singleton Park,
Swansea, SA2 8PP, UK.
e-mail: p.reed@swansea.ac.uk

Short title: Reinforcer devaluation.

Cite as: Reed, P. (in press). Reinforcer devaluation affects within-bout responses to a greater degree than bout-initiation responses on human free-operant responding. *Journal of Experimental Psychology: Animal Learning and Cognition*.

Abstract

Six experiments with human participants explored the effects of a reinforcer devaluation procedure on the microstructure of free-operant schedule performance, which is theoretically and empirically novel. The goal was to test whether different aspects of the microstructure of free-operant responding would be differentially affected by the manipulation to test the view that bout-initiation responses can be regarded as stimulus-driven habits both susceptible to devaluation, and within-bout responses as goal-directed actions that are susceptible to such effects. Humans responded on a fictious investment game in which responses made investments in two countries with different currencies. These currencies had an exchange rate into GBP, which could be manipulated to devalue the currency (reinforcer). In Experiments 1a and 2, RR schedules showed overall devaluation effects, which were limited to within-bout rates, and more pronounced on longer ratios. In Experiments 1b and 3, RI schedules showed limited devaluation effects, overall, except for with shorter intervals. This effect was seen on some schedules for within-bout responding. Experiments 4 and 5 manipulated the response-reinforcer feedback function (Experiment 4), and participants' ability to experience contingency variations, by pretraining to be variable or static in response rates (Experiment 5). Only when there was a stronger response-reinforcer relationship (Experiment 4), which could be experienced due to variable responding (Experiment 5), did the devaluation effect occur, and these effects were limited to within-bout responses. These data suggest that an account of free-operant responding based on there being, at least, two classes of response – bout initiations that are stimulus-driven habits, and within-bout responses that are goal-directed actions – may explain many schedule effects.

Keywords: Schedules of reinforcement; Microstructure; Reinforcer devaluation; RR schedule; RI schedule; Humans

It is suggested that responding on free-operant schedules of reinforcement comprises responses categorised as habits, which are not emitted directed at a goal; and those categorised as actions, which are goal-directed (Dickinson, 1985; Garr & Delamater, 2019; Perez & Dickinson, 2020). Several suggestions have been made about the conditions under which such habits and actions will come to dominate on particular reinforcement schedules (Chen & Reed, 2024; Dickinson, 1985; Perez & Dickinson, 2020; Thrailkill, Trask, Vidal, Alcalá, & Bouton, 2018). For example, Perez and Dickinson (2020; Dickinson, 1985) suggested that on schedules in which there is a strong relationship between variations in the rate of response and variations in the rate of received reinforcement, then responses will tend to be more sensitive to the goal and tend to be actions. This is the case on random ratio (RR) schedules, where there is a direct relationship between the rate of responding and the rate of reinforcement. In contrast, when there is not such a strong relationship between variations in rates of responding and reinforcement, such as on random interval (RI) schedules, habits come to be emitted in relatively greater numbers. Reinforcer devaluation studies have tended to support this view. Devaluation of reinforcers is taken to impact goal-directed actions, and such devaluation on RR schedules tend to have greater deleterious impacts on subsequent responding than devaluation on RI schedule peformance (Dickinson, Nicholas, & Adams, 1983; Pérez, Aitken, Zhukovsky, Soto, Urcelay, & Dickinson, 2019).

Recently, an alternative viewpoint has been suggested that, while partially consistent with the above view (Dickinson, 1985), seeks to develop and refine the suggestion (Chen & Reed, 2024; Reed, 2015; Reed, Smale, Owens, & Freegard, 2018). According to this newer view, responses on both RR and RI schedules can either be goal-directed actions or stimulus-driven habits; with the former being more prevalent on RR schedules, and the latter more prevalent on RI schedules (Reed et al., 2018). This account seeks to incorporate recent analysis of free-operant responding as being composed of two forms of responses controlled by different aspects of the contingency (Falligant, Hagopian, & Newland, 2024; Shull, 2011; Shull, Gaynor, & Grimes, 2001), and these analyses have been extended and replicated for human participants (Chen & Reed, 2024; Reed, 2015; Reed et al., 2018). Bout-initiation responses are taken to be largely controlled by factors such as the overall rates of reinforcement experienced in that context (independent of the nature of the schedule), and these responses have been taken to be stimulus-driven habits (Reed et al., 2018). Within-bout responses emitted after the first response of a bout, are taken to be goal-directed and controlled by reinforcement (Falligant et al., 2024; Reed & Chen, 2024).

Evidence for the view that bout-initiation responses can be equated with unconscious habits, and within-bout responses equated with conscious actions, comes from several sorts of study. Firstly, rates of bout-initiation responding vary with the rate at which reinforcement is delivered in that context for both nonhuman and human participants (Killeen, 2023; Reed, 2025; Reed et al., 2018; Shull, 2011), independent of the operative schedule, and are equivalent when rates of reinforcement are yoked across different schedules (Falligant et al., 2024; Reed et al., 2018). However, there are greater numbers of within-bout (action) responses on RR schedules (Reed et al., 2018; Shull, 2011). Secondly, manipulations that seek to make human participants more aware of the effective response-reinforcer relationship, such as giving verbal instructions, impact within-bout responding to a greater extent than bout-initiation responding, suggesting the latter is a largely unconscious stimulus-driven process (Chen & Reed, 2024). Thirdly, manipulations in humans that are taken to affect unconsciously driven processes and bring them into conscious awareness (like mindfulness), differentially reduce the level of bout-initiation responding in favour of within-bout responding (Chen & Reed, 2023; Reed, 2023). Finally, ERP studies have shown that patterns of brain activity associated with within-bout responses are typical of those associated with actions of which there is conscious awareness and control, but brain activity around bout-initiation responses is like that associated with stimulus-elicited responses (Chen, Shui, Chen, Zhang, & Reed, 2025).

Given the above, the suggestion is that all schedules support both habits and actions, which are identifiable with responses that initiate a bout (habits), or constitute part of a bout once initiated (actions). This is not at odds with previous suggestions that actions are more common on RR schedules (Dickinson, 1985), but offers clarification and specification on the responses that can be classed as habits and actions, as well as allowing some rapprochement between the literature derived from an associative tradition on habits and actions (e.g., Dickinson, 1985; Perez et al., 2019), and that largely derived from an operant tradition on the micro-structure of instrumental responding (e.g., Falligant et al., 2024; Reed, 2015).

In terms of the difference between RR and RI results after reinforcer devaluation studies, this new account suggests that, when the schedules are yoked to deliver the same amount of reinforcement in a particular context, then equal numbers of bout-initiation responses will be emitted. However, a far greater number of within-bout responses will be emitted on the RR schedule compared to an RI schedule. As bout-initiation responses are taken to be stimulus-controlled habits (Reed et al., 2018), reinforcer devaluation will not affect these responses as strongly as it affects within-bout responses that are taken to be goal-directed (see Dickinson et al., 1983; Perex & Dickinson, 2020). As overall responding on RR schedules has a greater proportion of within-bout responses than RI schedules, it will appear to be more greatly affected by reinforcer devaluation. Thus, there is no necessary contradiction between the effects to be expected based on the accounts, but the newer account gives greater specificity as to which responses should be impacted.

However, no study to date has directly assessed the effect of reinforcer devaluation on the micro-structure of free-operant responding in either humans or nonhumans. Given the clear predictions about the impact of reinforcer devaluation emerge from the view that actions and habits can be mapped onto bout-initiation and within-bout responding, the current experiments explored these effects. A version of the reinforcer devaluation paradigm developed by Allman, DeLeon, Cataldo, Holland, and Johnson (2010) for humans was utilised in the current context. In this paradigm, humans respond in the context of an investment game. A response is an investment in a country, which may or may not produce a financial return (reinforcer) in the currency of the country. The economy of the country can be manipulated so that it provides returns at different rates or according to different rules. The value of the currency used when converted to pounds sterling (dollars, etc.) can be reduced to allow for reinforcer devaluation manipulation. This procedure can be adapted to study the effect of devaluation on a variety of reinforcement schedules, and maps onto previous studies of human free-operant responding maintained in the context of an economic investment game (Reed, 2019). The first prediction is that reinforcer devaluation after exposure to an RR schedule will have a more noticeable effect on subsequent responding, than it does for RI schedule responding. The second prediction is that reinforcer devaluation will impact within-bout responding more readily than bout-initiation responding. Finally, it follows that, as there will be more within-bout responses on RR than on RI schedules, the effect though present on RI schedules, will be more noticeable on RR schedules.

Experiments 1a and 1b

Experiment 1 explored the effect of devaluing a reinforcer on responding maintained by either an RR schedule (Experiment 1a) or an RI schedule (Experiment 1b). Participants responded (pressed a computer key) in an investment game as previously used by Reed (1999). An initial investment fund of a fictitious currency was displayed on the screen. Each response made would reduce the total currency displayed by a set amount. Sometimes the investment (response) would gain a financial return (reinforcement) and add to the total amount of currency displayed. Reinforcement could be delivered according to an operative schedule (Reed, 1999). The reinforcer devaluation method was derived from that described by Allman et al. (2010). The financial returns obtained were given in a particular fictitious currency that could then the exchanged for the currency of the participant (Great British Pounds; GBP). The exchange rate could be reduced to devalue the reinforcer. In the first phase, participants were told that the exchange rate was 1:1, and then responded for some time at that rate. The subsequent devaluation involved informing participants that the value of the fictitious currency had collapsed making the exchange rate 1000:1. By using two fictitious countries, each with their own fictitious currency, presented separately in a multiple schedule (not concurrently; see Kosaki & Dickinson, 2010), one currency could be devalued and the other not devalued. Following this information being given, participants continued to respond to the two countries but without feedback (i.e. in extinction).

An RR schedule was used to relate responses to reinforcers (investments to returns) in the current Experiment 1a; whereas an RI schedule was used in Experiment 1b to organise the relationship between investments (responses) and returns (reinforcement). Previous findings have indicated that the best chance of observing a devaluation effect on an RI schedule is by using an interval value that produces a relatively high rate of reinforcement (e.g., Garr et al., 2020), and so an RI-15s schedule was employed for Experiment 1b. On the basis of previous studies, an RR schedule value (RR-25) that has produced a similar rate of reinforcement, using a similar procedure with humans, to the RI-15s schedule (Chen & Reed, 2022; Reed et al., 2018) was selected for Experiment 1a. The effects of devaluation on schedule performance in humans using this procedure have not been explored. However, should the effects reported by Garr et al. (2020) for rats on rich RI schedules be replicated, both RR and RI schedules should show devaluation effects.

In addition, the micro-structure of responding following devaluation was assessed. A 'cut-off' method was used to determine bout-initiation and within-bout responding (Mellgren & Elsmore, 1991; Reed et al., 2018; Sibley, Nott, *&* Fletcher, 1990). This was chosen in preference to the log survivor method (Shull, 2011), as this latter method has sometimes proved difficult to fit to observed responding (see Falligant et al., 2024; Shull et al., 2002). The cut-off method designates short inter-response times (IRTs) as within-bout responses, and long IRTs as bout-initiation responses (Mellgren & Elsmore, 1991). In the context of human schedule studies, an IRT of 1000ms has proved a good index of this distinction (Reed et al., 2018; Chen & Reed, 2022).

Bout-initiation responses are taken to be related to the status of the context; whereas within-bout initiation responses are affected by the status of the reinforcer (Falligant et al., 2024; Reed et al., 2018; Shull et al., 2002). On the basis of previous studies, testing in extinction should affect both forms of responding (Chen et al., 2020; Shull et al., 2002), and devaluation may impact both the status of the context and the reinforcer, which also may affect rates of bout-initiation (Bouton, 2021; Chevée, Kim, Crow, Follman, Leonard, & Calipari, 2023; Turner & Balliene, 2024). However, according to the view developed by Reed et al. (2018; see also Chen & Reed, 2024), devaluation should have a greater impact on goal-directed within-bout responses than on the stimulus-driven bout-initiation responses.

It may be that the sometimes-observed lack of devaluation effects seen on RI schedules (Chevee et al., 2023; Dickinson et al., 1983; Thrailkill & Bouton, 2015) are because responding on RI schedules is less goal-directed than that on RR schedules due to the weaker response-reinforcer relationship on interval compared to ratio schedules (Dickinson et al., 1983). However, the results presented by Garr et al. (2023) make this unlikely as the sole mechanism (see also Chevee et al., 2023; DeRusso et al., 2010). It may also be because a greater proportion of responses emitted on RI schedules, compared to RR schedules, are bout-initiation responses. As bout-initiation responses are dependent on obtained reinforcement rates in that context, their emission is not dependent on particular response-reinforcer schedules, and similar rates of reinforcement produce similar bout-initiation responding across different schedules (Chen & Reed, 2022; Reed et al., 2018; Shull, 2011). As overall rates of response are lower on RI schedules, then there will be fewer within-bout responses to be affected by the devaluation.

In sum, Experiments 1a and 1b explored whether devaluation effects would be seen on both RR and RI schedules; and whether they would be evident for both within-bout and bout-initiation responding. The schedules were studied in two separate experiments due to the problems that yoking procedures would produce for such a design. In both studies, participants were presented with a task in which they had to explore 'investment return' in two fictitious countries. Investments in the countries were made in the countries' (fictious) currencies, each of which had an exchange rate to Pounds Sterling. In Phase 1 the exchange rate was the same for each country, but in phase 2, participants told that one currency had collapsed, and the exchange rate was now very low (i.e. devaluation occurred). The currency in the other country was unaffected.

Method

Transparency and Openness.

We report how we determined sample size, data exclusions, and manipulations. Data are available at: https://osf.io/qvpab/overview. Data were analysed using SPSS version 26.

Participants

A sample of 55 students (27 for Experiment 1a; and 28 for Experiment 1b) were recruited via the Psychology Department subject-pool. G-Power calculations suggested that for a large effect size (*f'*=.35, based on findings reported by Allman et al, 2010), with 90% power, using a rejection criterion of $p<.05$, a sample size of 24 per study would be sufficient to detect an interaction in a two-factor repeated-measures analysis of variance (ANOVA). This is in line with the sample size reported by Allman et al. (2010). Participants received credits for their participation, and were also told that the person who had earned most hypothetical money during the experiment would receive a £100 book token (which was given).

Participants were aged between 18 to 25 years. No participant self-reported previous history of mental illness. However, as individuals high in depression or schizotypy can show atypical schedule performance (Dack, McHugh, & Reed, 2009; Randell, Ranjith-Kumar, Gupta, & Reed, 2009) they were screened for these traits. A score greater than 10 on Beck's Depression Inventory, and greater than 6 on the Unusual Experiences scale of the Oxford Liverpool Inventory of Feelings and Experiences Brief (one standard deviation above the mean), was taken as a cut-off for exclusion. Following psychometric measurement, two participants were excluded from Experiment 1a, and three from Experiment 1b, leaving 25 participants in each study. Ethical approval for this, and all studies reported here, was given by the Psychology Ethics Committee of Swansea University.

Apparatus

The experimental task was presented on a standard desktop computer. Visual Basic (6.0) was used to programme the task, which presented a multiple schedule. This was a multiple (mult) RR-25, RR-25 for Experiment 1a; and a mult RI-15s, RI-15s for Experiment 1b. Each component of the multiple schedules was presented until 10 investment returns (reinforcers) had been obtained, each followed by a 1s intercomponent interval. The components were differentially signalled by the words 'Country 1' or 'Country 2', appearing at the top of the monitor during exposure to the component. A picture representing the country, such as a desert scene or cityscape, was also presented as a background – the image-country pairing was counterbalanced across participants. A box was presented at the bottom the screen with a running total of ‘Amount earned’, which started at 40 units for each country on each exposure (the unit of currency was given a different name for each country, counterbalanced across participants). Information regarding the exchange rate for that country was presented at the top of the screen (e.g., 1 currency unit = 1 GBP; or 1 currency unit = .001 GBP). After the exposure to the schedule component, the display disappeared, and the participants were told how many GBP they had earned in the previous visit to the country, based on the current exchange rate, and this information was presented on a screen for 1s.

Experiment 1a used a mult RR-25, RR-25 schedule, for which each space bar response (investment), in each component (country), had a 1/25 probability of being followed by reinforcement (the addition of 40 currency units to the total). Experiment 1b used a mult RI-15s, RI-15s schedule, where each second during exposure to the RI components had an equal probability of being assigned as the period after which reinforcement would be delivered for a response (i.e. 1/15). Reinforcement (units of currency) was added for the first response made after that second. In both experiments, participants lost one currency unit for each space bar response, in each component, regardless of whether, or not, that response was reinforced. It has previously been established that a response cost generates schedule performance by humans similar to that observed in nonhumans, as the absence of response cost creates little reason to regulate performance (Bradshaw & Reed, 2012; Raia, Shillingford, Miller, & Baier, 2000).

Materials

Beck's Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) assesses depression (Cronbach α = .73 to .92: Beck, Steer, & Garbin, 1988). A score of 14 or higher is taken as a cut-off for individuals displaying levels of depression.

Oxford Liverpool Inventory of Feelings and Experiences - Brief Version (O-LIFE(B); Mason, Linney, & Claridge, 2005) measures schizotypy (Cronbach α = .62 to .80).

Procedure

Participants were tested individually in a quiet room, which contained a desk and computer, with the monitor situated approximately 60cm from them. Participants read the study information and instructions for the task. Participants commenced the task in their own time, and were required to fill in basic demographic details about themselves, and the BDI and O-LIFE questionnaires, before the schedule task was presented. When the questionnaires have been completed, participants were asked to start to study, and were directed to the instructions on the monitor:

"*There are two components to this task, which you have to do at the same time as each other. For the first, you will be given a number, and will have to count outload backwards in 7s. Your performance will be recorded for later accuracy analysis. For the other task, you have to use the space bar to invest in the economy of a country in order to earn hypothetical money. You will be exposed to two different countries, each country has its own currency. From time to time, you will be shown the exchange rate between that country's current and GPB, this will remain constant until you are told otherwise, but the exchange rate may vary from time to time. There will be several exposures to each of these two different countries. The different countries will be signalled 'Country 1' or Country 2' at the top of the screen. Exposure to the countries will alternate. Your goal during each exposure* *to a country is to earn as much money as possible. Sometimes it may be best to invest at different rates.*".

In each experiment reported in this series, participants were exposed to two components (countries), one which would be become the devalued conditioned (country), and one which would be the non-devalued or control condition (country). Of course, it would have been possible to have more countries, but this would have increased the length of the experiment. Each schedule presentation (trial) lasted until 10 outcomes had been obtained. In phase 1, there were eight such presentations of schedule components (RR in Experiment 1a, and RI in Experiment 1b); four labelled 'Country 1', and four labelled 'Country 2', which alternated. At the start of each exposure to each country in Phase 1, an exchange rate of 1:1 unit:GPB was given. The participants were then given information that the currency in one of the countries (counterbalanced for participants) had suffered an economic collapse, as was now only worth 1/1000 of a GBP. The currency in the other country was unaffected, and still had an exchange rate of 1:1 with GBP ("*There has been an economic cllapse in Country 2, and its current is now worth only 1/1000 of a GBP. The exchange rate for Country 1 is unaffected.*"). The county that suffered the collapse was counterbalanced across participants. Phase 2 followed, and comprised a further 8 exposures to the components (4 for each country), but under conditions of no reinforcement (extinction). Each of the four exposures to each country lasted for 60s.

Throughout exposure to the schedules, participants had to perform a counting backwards task throughout the entire experiment (Andersson, Hagman, Talianzadeh, Svedberg, & Larsen, 2002). They were each given one random five-digit number at the start of the procedure (different for each participant), and were asked to count backwards from that number, out-loud, in 7s. This procedure was adopted to minimise verbal rule formation which can influence participants' performance on schedules (Bradshaw & Reed, 2012;

Leander et al., 1968; Raia et al., 2000). To enhance task adherence, a recording device was placed prominently on the desk in front of the participant, and they were told that their answers to the counting task would be analysed and scored later (this analysis was not performed).

Results and Discussions

Experiment 1a

Figure 1 about here

Figure 1 (top panel) shows overall response rates for the last exposure to each RR-25 component (country) during phase 1 (before devaluation), and also the rates averaged across the extinction test following devaluation in Phase 2. By the end of Phase 1, there were similar overall rates of response in both countries (mean reinforcement rate = 4.55$\pm$2.04SD reinforcers/min). However, in Phase 2 (following devaluation) during extinction, overall response rates became lower for both components, but more so for the devalued component. A two-factor repeated-measures ANOVA, with component (devalued x control) and phase as factors, was conducted on these data. There were significant main effects of phase, $F(1,24)=119.17$, $p<.001$, MSe=907.00, $\eta^2_p=.832$[95%CI =.019:,487], and component, $F(1,24)=9.56$, $p=.005$, MSe=57.25, $\eta^2_p=.285$[.019:,487], and a significant interaction, $F(1,24)=36.09$, MSe=51.84, $p<.001$ $\eta^2_p=.601$[.004:,445]. Simple effect analyses revealed no difference between components in Phase 1, $p>.05$, but a significant difference in Phase 2 with lower overall rates in the devalued component, $F(1,24)=42.71$, $p<.001$, MSe=51.84, $\eta^2_p=.267$[.032:,512].

The middle panel of Figure 1 displays the bout-initiation rates, using a 1000ms cut-off analysis (Reed et al., 2018), and shows a decrease in responding during extinction for both components, but no clear differential impact of devaluation. A two-factor repeated-measures ANOVA (component x phase) revealed a main effect of phase, $F(1,24)=108.38$, $p<.001$, MSe=59.58, $\eta^2_p=.819[.004:,445]$, but no main effect of component (devaluation), nor interaction, both *p*s>.08.

The bottom panel of Figure 1 shows the within-bout rates, which were similar for both components at the end of phase 1. These decreased in Phase 2, but more markedly for the devalued component. A two-factor ANOVA (component x phase) revealed a significant main effect of phase, $F(1,24)=88.66$, $p<.001$, MSe=1269.42, $\eta^2_p=.787[.026:,501]$, no main effect of component, $F<1$, but a significant interaction, $F(1,24)=30.02$, , $p<.001$, MSe=270.74, $\eta^2_p=.556[.006:,451]$. Simple effect analyses revealed a higher within-bout rate for the to-be-devalued component in phase 1, $F(1,24)=16.55$, $p<.001$, MSe=270.74, $\eta^2_p=.294[.036:.518]$, but a there was a significantly lower rate for the devalued than for the non-devalued component in phase 2, $F(1,24)=13.62$, $p<.001$, MSe=270.74, $\eta^2_p=.294[.036:,518]$.

These findings confirm that the current procedure was effective in reducing responding to a devalued component of an RR-25 schedule. That devaluation effects were readily apparent on an RR schedule is in line with previous demonstrations with nonhumans (Dickinson et al., 1983). Analysis of the micro-structure of responding demonstrated that within-bout rates of responding had been impacted by the devaluation manipulation, but that bout-initiation responses were not impacted. This is in line with suggestions made by Reed et al. (2018), and Chen and Reed (2024), that within-bout responding is goal-directed, but bout-initiation responding is largely stimulus-driven.

Experiment 1b

Figure 2 about here

Figure 2 (top panel) shows the overall response rates for the last exposure during each phase for both RI components. There were similar overall rates of response in both components at the end of Phase 1. Response rates decreased in both components during Phase 2 (extinction), but decreased more in the devalued component. A two-factor ANOVA (component x phase) revealed significant main effects of phase, $F(1,24)=20.27$, $p<.001$, MSe=1379.01, $\eta^2_p=.458$[.004:.445], and component, $F(1,24)=6.76$, $p=.016$, MSe=162.15, $\eta^2_p=.220$[.004:.445], and an interaction, $F(1,24)=4.77$, $p=.041$, MSe=216.79, $\eta^2_p=.163$. Simple effect analyses revealed no difference between components in Phase 1, $F<1$, but a significant difference in Phase 2 with the devalued component having a lower rate, $F(1,24)=9.75$, $p<.001$, MSe=216.79, $\eta^2_p=.267$[.032:,512].

The middle panel of Figure 2 displays the bout-initiation rates, using a 1000ms cut-off analysis. A two-factor ANOVA (component x phase) a significant main effect of phase, $F(1,24)=9.03$, $p=.006$, MSe=10.31, $\eta^2_p=.273$, but no other main effect nor interaction, both Fs<1. The bottom panel of Figure 2 shows that by the end of Phase 1 there were similar within-bout rates in both components. Within-bout rates decreased for both components during Phase 2, but more so for the devalued component. A two-factor ANOVA (group x phase) revealed a significant main effect of phase, $F(1,24)=16.51$, $p<.001$, MSe=1482.22, $\eta^2_p=.408$[.003:.444], no main effect of component, $p>.08$, but there was an interaction, $F(1,24)=5.99$, $p=.022$, MSe=224.42, $\eta^2_p=.200$. Simple effect analyses revealed no difference between components in Phase 1, $F<1$, but a significant difference in Phase 2 with lower rates in the devalued component, $F(1,24)=6.44$, $p=.018$, MSe=224.42, $\eta^2_p=.267$[.032:,512].

The data from Experiment 1b using an RI schedule show a devaluation effect on overall response rates. This effect is in line with previous demonstrations of devaluation effects after exposure to RI schedules using nonhumans with similar rich RI schedules (Garr et al., 2020). The devaluation effected within-bout responding to a much greater extent than bout-initiation responding. These data suggest that actions (within-bout responses) and habits (bout-initiation responses) are both present on RI schedules.

Experiment 2

Experiment 2 sought to confirm the results of Experiment 1a with respect to RR schedules, and to extend the study to a wider range of RR schedule values. On the basis of the results of Experiment 1, the devaluation effect was expected to be present, but to be greater for within-bout responses, than for bout-initiation responses on these schedules.

Method

Participants and Apparatus

A sample of 128 students, between 18 and 25 years old, was recruited via the Psychology Department subject-pool, as described in Experiment 1. After psychometric measurement, 8 participants were excluded on this basis, leaving 120 in the study. The apparatus and materials were as described in Experiment 1

Procedure

The procedure was as described in Experiment 1, with the following exceptions. There were three groups of participants (n=40): one responded on an RR-10 schedule in each component of the multiple RR, RR schedule; one group on a mult RR-30, RR-30 schedule;

and one on a mult RR-60, RR-60 schedule. For each group, Phase 1 had 8 alternating presentations of each RR schedule (four ‘Country 1’, and ‘Country 2’), each lasting until 10 reinforcers were delivered. At the start of each exposures, an exchange rate of 1:1 unit:GPB was given. Phase 2 was preceded by a ‘financial update’ informing participants that one of the countries had experienced an economic collapse, as described in Experiment 1. There followed, a further 8, 60s exposures to the components in extinction (4 for each country).

Results and Discussion

Figure 3 about here

The top panel of Figure 3 shows the group-mean overall response rates for the devalued and control components, for the groups (RR-10, RR-30, and RR-60), on the last block of Phase 1 and during Phase 2 (extinction test) training. These data show that there was an increase in response rate from RR-10 to the RR-30 schedule, which decreased a little for the RR-60 schedule. The rate of responding on these schedules was largely independent of the overall rate of reinforcement obtained (RR-10 = 8.61$\pm$2.46 reinforcers/min; RR-30 = 3.97$\pm$1.04; RR-60 = 1.78$\pm$.67). This biotonic relationship has been noted several times previously as ratio sizes increase across this range of values (Crossman, Bonhem, & Phelps, 1987; Reed & Hall, 1989). Rates of response decreased in the devalued component in Phase 2 for all schedules, confirming and extending the finding for RR-25 in Experiment 1 (see also Dickinson et al., 1983).

A three-factor mixed-model ANOVA, with group (RR-10, RR-30, RR-60) as a between-subject factor, and component (devalued v control) and phase as within-subject factors, revealed main effects of group, $F(2,117)=7.84$, $p<.001$, MSe=2233.94,

η^2_p=.118[.206:.681], phase, $F(1,117)$=1326.51, p<.001, MSe=351.91, η^2_p=.919[.206:.681], and component, $F(1,117)$=3.91, p=.050, MSe=2052.46, η^2_p=.032[.206:.681]. There were two-way interactions between group and phase, $F(2,117)$=18.07, p<.001, MSe=351.91, η^2_p=.236[.206:.681], and phase and component, $F(1,117)$=14.36, p<.001, MSe=306.92, η^2_p=.109[.206:.681], but not between group and component, nor between all three factors, both Fs< 1. Simple effect analysis revealed no difference in the rate for the control and devalued components in Phase 1, F<1, but revealed a significantly lower rate for the devalued component in Phase 2, $F(1,117)$=39.78, p<.001, MSe=306.92, η^2_p=.150[.206:.681].

The middle panel of Figure 3 displays the bout-initiation rates, which show that the RR-10 group had higher bout-initiation rates that the other two groups, and that responding decreased across the phases. This is consistent with bout-initiation responding being emitted dependent on the rate at which reinforcement is obtained in that context (Chen et al., 2020; Shull, 2002). They also show that devaluation had little differential effect on the two countries. These observations were corroborated by a three-factor ANOVA (group x phase x country) which revealed significant main effects of group, $F(1,117)$=194.05, p<.001, MSe=5.59, η^2_p=.768[.206:.681], and phase, $F(1,117)$=745.14, p<.001, MSe=1.10, η^2_p=.864[.206:.681], and an interaction between phase and group, $F(1,117)$=122.31, p<.001, MSe=1.10, η^2_p=.864[.206:.681], but no other effects or interactions, including any main effects or interactions involving component (devaluation), were significant, all ps > .06.

The bottom panel of Figure 3 shows that within-bout rates for the groups across the phases, which followed a similar pattern to the overall rates, with responding decreasing in the devalued component in Phase 2 for all schedules. A three-factor ANOVA (group x phase x country) revealed significant main effects of phase, $F(1,117)$=817.81, p<.001, MSe=985.38, η^2_p=.875[.206:.681], a significant phase and group interaction, $F(2,117)$=5.86, p=.004, MSe=985.38, η^2_p=.091[.206:.681], and a significant phase and component interaction, *F*(2,117)=3.89, *p*=.005, MSe=990.03, η^2_p=.033[.206:.681]. No other main effect or interactions were significant, all *p*s > .07. Simple effect analysis comparing the Phase 1 rates found no significant difference between control and devalued components, *F*<1, but there was a significantly lower rate in the devalued component in Phase 2, *F*(1,117)=20.06, *p*<.001, MSe=990.03, η^2_p=.150[.206:.681].

The analysis of the micro-structure reveals that devaluation impacted within-bout responding to a greater extent than it did bout-initiation responding across the RR schedules. This confirms for a greater range of RR values what was found in Experiment 1a for an RR-25 schedule. It also suggests that the analysis of the differential control of bout-initiation and within-bout responding developed previously (Reed et al., 2018) is broadly correct in that within-bout to a greater extent than bout-initiation responding are goal-directed and sensitive to alterations in the value of the outcome.

Experiment 3

Experiment 3 extended the range of RI schedules studied from that in Experiment 1b (RI-10, RI-30, RI-60). This would allow exploration of whether the effect was dependent on the schedule parameters. Garr et al. (2020) suggested that the interval schedule value was a determinant of when actions and habits emerged on RI schedules, with actions emerging earlier for richer schedules than for leaner RI schedules. If this were the case, then the devaluation effect might be expected to be more readily apparent on RI-10s and RI-30s schedules than on RI-60s schedule, especially with the relatively short training given in this procedure. Garr et al. (2020) suggested that this may be due to a greater correlation between response and reinforcement rates experienced on richer RI schedules, which is an argument that has also been advanced by Baum (1993). Studying a range of RI values would also allow comparison of RI schedules over a similar range of reinforcement rates that overlapped with those studied in Experiments 2 for RR schedules.

Method

Participants and Apparatus

A sample of 125 students, between 18 and 25 years old, was recruited via the Psychology Department subject-pool, as described in Experiment 1. Participants received credits for their participation, but no financial payment. After psychometric measurement, 5 participants were excluded on this basis, leaving 120 in the study. The apparatus and materials were as described in Experiment 1

Procedure

The procedure was as described in Experiment 2, with the exception that three groups of participants (*n*=40) responded on a mult RI RI schedule: one group on a mult RI-10s RI-10s schedule; one on a mult RI-30s RI-30s schedule; and one on a mult RI-60s RI-60s schedule. In Phase 1, there were 8 alternating presentations of each RI schedule (four ‘Country 1’, and four ‘Country 2’), each lasting until 10 reinforcers had been delivered. At the start of each exposure, an exchange rate of 1:1 unit:GPB was given. Phase 2 followed instructions that the exchange rate of one of the countries had collapsed, as described in previous studies reported here. Phase 2 comprised a further 8, 60s exposures to the components, but this time in extinction (4 for each country).

Results and Discussion

Figure 4 about here

Figure 4 (top panel) shows the overall response rates for the three groups, for both components of the mult RI, RI schedule, on the last trial of Phase 1 and averaged across trials of Phase 2 (extinction test). Inspection of these data shows that response rates were higher for the RI-10s group, and lowest for the RR-60s group, which is in line with the rates of reinforcement experienced by the groups. For all groups, rates of responding decreased during extinction (Phase 2). The rate decreased more for the devalued than the control component in Phase 2 for the RR-10s and RI-30s schedule groups, but less so for the RI-60s group. A three-factor mixed-model ANOVA (group x phase x component) revealed significant main effects of group, $F(2,117)=59.35$, $p<.001$, MSe=245.43, $\eta^2_p=.504[.003:.444]$, and phase, $F(1,117)=994.74$, $p<.001$, MSe=32.62, $\eta^2_p=.895[.003:.444]$. There were significant interactions between group and phase, $F(2,117)=44.95$, $p<.001$, MSe=32.62, $\eta^2_p=.435[.003:.444]$, phase and component, $F(2,117)=42.50$, $p<.001$, MSe=48.95, $\eta^2_p=.266[.003:.444]$, and all three factors, $F(2,117)=4.03$, $p=.020$, MSe=48.95, $\eta^2_p=.064[.003:.444]$. No other main effects nor interactions were significant, all ps>.20. Simple effect analysis comparing the Phase 1 rates for the control and devalued components revealed no differences for any group, all ps>.10. Simple effect analyses noted a significantly lower rate for the devalued component for the RI-10s and RI-30s schedules, smallest $F(1,117)=5.78$, $p<.001$, MSe=171.45, $\eta^2_p=.150[.206:.681]$, but not for the RI-60s schedule, $F<1$.

The middle panel of Figure 4 shows the bout-initiation rates generally decreased as the size of the RI schedule increased (they were highest in the RI-10s group), which is in line with the suggestion that bout-initiation responding is driven by the value of the context (Shull et al., 2002). There was no effect of devaluation on these responses. A three-factor mixed-model ANOVA (group x phase x component) revealed significant main effects of group, $F(2,117)=33.25$, $p<.001$, MSe=11.64, $\eta^2_p=.362[.003{:}.444]$, and phase, $F(1,117)=259.59$, $p<.001$, MSe=4.28, $\eta^2_p=.689[.003{:}.444]$, and an interaction between group and phase, $F(2,117)=10.80$, $p<.001$, MSe=4.28, $\eta^2_p=.156[.003{:}.444]$, but none of the other main effects nor interactions (i.e. no effect/interaction involving component/devaluation) were significant, all *p*s>.06.

The bottom panel of Figure 4 shows the within-bout rates for the three groups across the components and phases. These data follow a similar pattern to those for the overall rates, and show a numerical devaluation effect for all groups (schedule types). A three-factor mixed-model ANOVA (group x phase x component) revealed significant main effects of group, $F(2,117)=17.10$, $p<.001$, MSe=2681.94, $\eta^2_p=.226[.003{:}.444]$, and phase, $F(1,117)=39.99$, $p<.001$, MSe=1472.18, $\eta^2_p=.255[.003{:}.444]$. There was a significant two-way interaction between component and phase, $F(2,117)=3.98$, $p=.048$, MSe=1446.86, $\eta^2_p=.033[.003{:}.444]$, but none of the other main effects nor interactions were significant, all *p*s>.08. Simple effect analysis comparing the rate in the control and devalued components revealed no significant effect in phase 1, $F<1$, but a significantly lower rate in the devalued component in phase 2, $F(1,117)=7.99$, $p=.048$, MSe=1446.86, $\eta^2_p=.033[.003{:}.444]$

These data imply that a devaluation effect can be obtained for RI schedules, albeit of a smaller magnitude than that seen for RR schedules (cf. Experiments 1 and 2). The effect is manifest mainly in the reduction of within-bout responding, and not for bout-initiation responding. As bout-initiation responses are a greater proportion of responses for RI than for RR schedules, this may play a role in the comparatively smaller effects seen, and contribute to the difficulties sometimes previously noted in obtaining this effect (Dickinson et al., 1983;

present Experiment 1b). The devaluation effect was only present at the richer RI values when considering overall responses rates, which corroborates the findings reported by Garr et al. (2020). However, when considering within-bout rates the effect was present in all schedules (albeit numerically greater for the richer schedules). This may be, again, as bout-initiation responding comprised a greater proportion of total responding for the leaner schedules in extinction. As devaluation does not seem to impact these responses so readily, then this may mask any effect that is present on leaner schedules with little training.

Experiment 4

The preceding experiments have verified several findings that have emerged from the nonhuman conditioning literature with respect to the impact of reinforcer evaluation on free-operant responding. They have noted that a devaluation effect is readily present on RR schedules (Dickinson et al., 1983), and can, under some conditions, be seen on RI schedules (Chevee et al., 2023; DeRusso et al., 2010; Garr et al., 2023). However, the current experiments also found that the devaluation effect was typically confined to within-bout responses. This limited devaluation effect is consistent with the analysis of human free-operant responding given by Reed et al. (2018; see also Chen & Reed, 2022) which suggests within-bout responses are largely goal-directed and under the control of the reinforcer; whereas, bout-initiation responses are stimulus-driven and under the control of the context. While the factors that are responsible for bout-initiation responding seem relatively clear, and are associated with the conditioned value of the context (Falligant et al., 2024; Reed et al., 2018), as are habits (Killeen, 2023), the precise factors responsible for generating goal-directed responding (within-bout responses) are less clear.

The current set of studies, coupled with recent previous work (Chevee et al., 2023; Garr et al., 2020), suggests that the occurrence of the devaluation effect is not just dependent on the nature of the schedule (e.g., RR versus RI). Although the rate of reinforcement appears to be important in this regard (Baum, 1993; Garr et al., 2020), this cannot be the full explanation as both Experiment 2 and Experiment 3 noted a devaluation effect for within-bout responding across a range of reinforcement rate values.

It has been suggested that the correlation between response rate and reinforcement rate is a key driver of the development of goal-directed actions (Dickinson et al., 1983; Garr et al., 2020). The response-reinforcer rate relationship has been shown to play a role through correlational analysis of the association between these variables on RI schedules, reported by Garr et al. (2020). This analysis suggests that it is the organisation of the relationship between responding and reinforcement that may be critical. The post hoc analyses provided by Garr et al. (2022) were indicative of the importance of this effect, but there are other methods of exploring this relationship.

Several schedules have been developed that will have the features of an RI schedule, but which will organise a positive relationship between response rate and reinforcement rate, and these schedules are sometimes referred to a random interval with linear feedback loop (RI+; McDowell & Wixted, 1988). The RI+ schedule is important as it allows direct assay of the importance of response-reinforcer rate associations. On an RI+ schedule, overall reinforcement rate depends on overall response rate, whereas this is not true across all the range of response rate values for a standard RI schedule. On an RI+ schedule, the interval required before reinforcement becomes available is given by: $a = (i / n) * b$; where, a = interval required before a response will produce reinforcement; i = time from last reinforcement to present; n = number of responses made during i; and b = equivalent RR value. For example, if the RI+ schedule were to have the feedback characteristics of an RR-

30 schedule, the time from the last reinforcer was 60s, and 60 responses had been made during that period; then the interval to next reinforcement would be (60/60)*30 = 30s. During this 30s, at 60 responses per min, 30 responses would be emitted before the interval times out. Alternatively, if 30 responses had been made during the 60 s period, then the interval would be: (60/30)*30 = 60s. During this 60s, if responses were emitted at 30 responses per minute, then 30 responses would be emitted prior to the reinforcer. Thus, a positive relationship exists between responding and reinforcement in a way it does not for an RI schedule.

The current study examined the effect of devaluation on an RI+ schedule, along with an equivalent RR and reinforcement-rate matched RI schedule, as well as examining the microstructure of responding in all cases. If Garr et al. (2020) are correct, then it might be expected that a devaluation effect would be seen on the RR and RI+ schedules at values that typically do not produce an effect on an RI schedule (e.g., RR-60 which would produce a relatively leaner rate of reinforcement than smaller ratio requirements).

Method

Participants and Apparatus

A sample of 167 students, between 18 and 25 years old, was recruited via the Psychology Department subject-pool, as described in Experiment 1. After psychometric measurement, 17 participants were excluded on this basis, leaving 150 in the study. The apparatus and materials were as described in Experiment 1

Procedure

The procedure was as described in Experiment 1, with the following exceptions. There were three groups of participants (*n*=50): one group responded on a mult RR-60, RR-

60 schedule; one on a mult RI-45s, RI-45s schedule (this interval value was chosen on the basis of previous studies (Chen & Reed, 2024; Reed et al., 2018) that have shown this is approximately the rate of reinforcement obtained on an RR-60 schedule); and one group on a mult RI+60, RI+60 schedule.

The RI+ schedule was programmed so the feedback function relating response to reinforcer rate was the same as a RR-60 schedule. That is, the probability of reinforcer occurring for a response in any given second was determined by the function: $a = (i / n) * b$; where: i = time from last reinforcement to present; n = number of responses made during i; and b = equivalent RR value, making a the value that determined the probability that a response in that period would be reinforced (i.e. 1/a).

For each group, Phase 1 had 8 alternating presentations of each RR schedule (four labelled 'Country 1', and four labelled 'Country 2'), each lasting until 10 reinforcers had been delivered. At the start of each of these exposures, an exchange rate of 1:1 unit:GPB was given. Phase 2 followed a financial update, as in Experiment 1, that the economy of one of the countries had collapsed. This extinction phase comprised a further 8, 60s exposures to the components (4 for each country).

Results and Discussion

Figure 5 about here

The top panel of Figure 5 shows the overall response rates for the three groups, for both components of the mult schedule, on the last trial of Phase 1, and averaged across the trials of the extinction test (Phase 2). Inspection of these data shows that response rates were higher for the RR and RI+ schedules than for the RI schedule. The rates of reinforcement for the schedules (averaged across both components were: RR = 1.7±.5; RI = 1.4±.4; RI+ = 1.6±.4). For all groups, rates of responding decreased during extinction (Phase 2), but the difference between devalued than the control components was greater for the RR schedule, and was in the same direction but of smaller size on the RI+ schedule. A three-factor mixed-model ANOVA (group x phase x component) revealed that all main effects and interactions were significant, all *p*s<.001, including the three-way interaction, $F(2,147)=40.95$, $p<.001$, MSe=16.17, $\eta^2_p=.358[.003:.444]$. Simple effect analysis comparing the Phase 1 rates for the control and devalued components revealed no differences for any group, all *F*s<1. For phase 2 (extinction), there was a significantly lower rate for the devalued component for the RR schedule, $F(1,147)=10.09$, $p<.001$, MSe=870, $\eta^2_p=.150[.206:.681]$, but not for the RI or RI+ schedule, both *ps*>.10.

The middle panel of Figure 5 shows the bout-initiation rates, which were lower during extinction (Phase 2) than in training (Phase 1) for all schedules. This would be expected given the largely similar rates of reinforcement experienced on the schedules (Shull, 2011). There was no effect of devaluation on these responses. A three-factor mixed-model ANOVA (group x phase x component) revealed a significant main effect of phase, $F(2,147)=191.93$, $p<.001$, MSe=3.03, $\eta^2_p=.566[.003:.444]$, but no other main effect nor interaction was significant corroborating that devaluation (component) had no impact on this responding, all *p*s> 0.10.

The bottom panel of Figure 5 shows the within-bout rates for the three groups across the components and phases. These data follow a similar pattern to those for the overall rates, and show a devaluation effect for the RR and RI+ schedules. A three-factor mixed-model ANOVA (group x phase x component) revealed all main effects and interactions to be significant, all *p*s<.001, including the three-way interaction, $F(2,147)=11.29$, $p<.001$, MSe=61.891, $\eta^2_p=.133[.003:.444]$. Simple effect analysis comparing the rate in the control and devalued components in Phase 1 revealed no significant effects, $F<1$. There was a significantly lower rate in the devalued component for the RR schedule in Phase 2, $F(1,147)=12.49$, $p<.001$, MSe=1124, $\eta^2_p=.033[.003:.444]$, as there was for the devalued RI+ schedule, $F(1,147)=5.13$, $p=.024$, MSe=1124, $\eta^2_p=.033[.003:.444]$, but not for the RI schedule, $F<1$.

The current results suggest that placing a linear feedback loop onto an RI schedule makes performance on that schedule more similar to that seen on an RR schedule. This finding has been noted several times previously for humans at the level of overall behaviour (McDowell & Wixted, 1988), and the microstructure of behaviour (Reed et al., 2018). Additionally, the current study also demonstrated that an RI+ schedule demonstrates similar devaluation effects to an RR schedule, and different to those of an RI schedule. This corroborates the hypothesis developed by Garr et al. (2020) that the correlation between response rate and reinforcement rate is important in determining when goal-directed actions will develop relatively quickly. However, it should be noted that, although the results from the RI+ schedule were different from those of the RR schedule, they did not fully match those of the RR schedule, in that devaluation effects only affected within-bout rates for the RI+ schedule, and not overall rates.

Experiment 5

That the correspondence between devaluation effects on RR and RI+ schedules in Experiment 4 was not complete implies that further factors than the response-reinforcer correlation are important in determining the development of goal-directed actions. In previous work exploring the impact of various response-reinforcer correlations on responding, it has been noted that the scheduled association between these aspects of free-operant responding is important, but that it will only play a role if it is experienced (Baum, 1993; Garr et al., 2020). This means that not only must there be such a programmed contingency, but that the behaviour must be variable enough for this association to be sampled (Reed, 2007). That is, there must be both periods of slower and quicker responding, so that the relationship between these response rates and the rate at which reinforcement is delivered can come to play a role.

There are various methods that have been adopted to encourage or decrease variability in rates of responding on free-operant schedules, often by the addition of a conjoint schedule component that reinforces variable or static responding (Page & Neuringer, 1985). A difficulty with these methods is that the additional conjoint component impacts rates of reinforcement, which are very difficult to control, and, in the present procedure, it is far from clear how this equation of reinforcement rate could be achieved. Another set of methods has also been attempted, which involves pretraining variable or static responding prior to exposure to the target contingency, in the hope that the pattern of responding will continue after training (Reed, 2007). These procedures have the advantage of not interfering with rates of reinforcement during exposure to the target schedule, but have the disadvantage that their effects on response variability will eventually be overridden by the operative schedule. However, the difficulties of managing the rate of reinforcement engendered by the first set of procedure suggested that similar approach to the problem may be indicated for the current study.

To this end, groups of participants responded on an RR-60, RI+60, or an RI-40s schedule (the value chosen based on rates of reinforcement obtained in Experiment 3), as well as two further groups RI+Var and RI+Static. These groups had either been pretrained on a schedule that delivered reinforcement if local response rates fell outside (Var), or within (Static), the rates previously emitted (Reed, 2007). The schedule would potentially reinforce responding on an RI-10s schedule. However, delivery of this reinforcement was dependent upon the response rate emitted during the preceding 10s either being: 5 responses per min over or under the rate emitted in the 10s period prior to that (Var); or within 5 responses per min of the rate emitted in the previous 10s period (Static). Thus, rates of response over successive 10s intervals were compared across the session, and if the local response rate in the preceding 10s (interval n) was over/under or within 5 responses per minute to that in the 10s preceding that interval (i.e., interval n-1), then the first response following the elapse of the 10s would be reinforced.

It was expected that devaluation effects would impact responding on RR and RI+Var schedules similarly; and that devaluation would impact RI and RI+Static responding similarly; with RI+ responding being somewhere between the two extremes. Such a pattern of results would corroborate the suggestions made by Garr et al. (2020), but with the addition of a provision that the correlation must be experienced in behaviour. The effect on the micro-structure is also expected to show much greater impacts on within-bout than bout-initiation responding.

Method

Participants and Apparatus

A sample of 217 students, between 18 and 25 years old, was recruited via the Psychology Department subject-pool, as described in Experiment 1. After psychometric measurement, 17 participants were excluded on this basis, leaving 200 in the study. The apparatus and materials were as described in Experiment 1

Procedure

The procedure was as described in Experiment 3, with the following exceptions. There were five groups of participants (*n*=40): one would respond on an RR-60 schedule in each component of the multiple RR, RR schedule; one on a mult RI-40s, RI-40s schedule; and three on a mult RI+60, RI+60 schedule, as described in Experiment 4.

An initial stage of training was included for this experiment, in which participants were exposed to a multiple schedule. The two components where labelled 'Country X' and 'Country Y', and each received reinforcement (investment returns) in their own units of currency. There were 6 exposures to each component (12 exposures in total), each lasting until 8 outcomes had been collected. Each component, within each group, had identical contingencies to one another. For one group (RI+Var), both components delivered reinforcement (investment returns) on an RI-15s schedule, if the rate in the current 10s interval was either 5 responses a min higher or lower than in the 15s period before. A second group (RI+Static) received these contingencies but with the difference that reinforcement was delivered if responding in the current 10s interval was within 5 responses a min of the preceding 10s period. The other three groups received mult RI-30s RI-30s schedules (based on the mean rate of reinforcement experienced in the other two groups during this period).

Following this training, the procedure was as described in Experiment 1, with the following exceptions. One group responded on a multiple RR-60, RR-60 schedule; one on a mult RI-40s, RI-40s schedule (this interval value was chosen on the basis of previous studies which have shown this is approximately the rate of reinforcement obtained on an RR-60 schedule; Chen & Reed, 2024; Reed et al., 2018); and three groups on a mult RI+60, RI+60 schedule, as described in Experiment 4. For each group, Phase 1 had 8 alternating presentations of each RR schedule (four labelled 'Country 1', and four labelled 'Country 2'), each lasting until 10 reinforcers had been delivered. At the start of each of these exposures, an exchange rate of 1:1 unit:GPB was given. Phase 2 followed a financial update, as described in Experiment 1, and comprised a further 8, 60s exposures to the components, but this time in extinction (4 for each country).

Results and Discussion

To ensure that the pretraining manipulation had been effective, the speed of earning the 8 reinforcers, averaged across the final exposure to both components, was calculated. The mean time to collect the 8 reinforcers for participants who were to be in the groups were: RR = 133.41$\pm$6.84s; RI = 131.57$\pm$5.05s; RI+ = 131.12$\pm$4.75; RI+Var = 132$\pm$5.80; RI+Static = 132.04$\pm$6.37. There were no significant differences in these data, $F<1$, suggesting the pretraining contingencies had been mastered, and participants who had a variability contingency were responding appropriately variably, and those who had a static component were responding appropriately less variably.

Figure 6 about here

The overall response rates for the three groups, for both components, on the last trial of phase 1, and averaged during extinction (Phase 2), are shown in the top panel of Figure 6. These data shows that response rates were higher for the RR, RI+, and RI+Var schedules than for the RI schedule. There was a slightly less pronounced difference between the RI+Static and RI schedules. Rates of reinforcement for the schedules (averaged across both components were: RR = 1.7$\pm$.6; RI = 1.5$\pm$.2; RI+ = 1.7$\pm$.5; RI+Var = 1.8$\pm$.6; RI+Static = 1.5$\pm$.4). Rates of responding decreased from these levels during Phase 2 (extinction). The difference between the devalued and control components was greater for RR and RI+Var schedules. A three-factor mixed-model ANOVA (group x phase x component) revealed that all main effects and interactions were significant, all *p*s<.001, including the three-way interaction, $F(4,195)=6.27$, $p<.001$, MSe=375.14, $\eta^2_p=.114$[.002:.364]. Simple effect analysis comparing the Phase 1 rates for the control and devalued components revealed no differences for any group, all *F*s<1. For Phase 2 (extinction), there was a significantly lower rate for the devalued component for the RR schedule, $F(1,95)=70.01$, $p<.001$, MSe=375.14, $\eta^2_p=.646$[.396:.981], and the RI+Var schedule, $F(1,95)=22.94$, $p<.001$, MSe=375.14, $\eta^2_p=.345$[.104:.597], but not for the RI ($p=.063$), RI+ ($F<1$), or RI+Static ($p=.065$) schedules.

Group-mean bout-initiation rates are shown in the middle panel of Figure 6. These were lower during extinction (Phase 2) than in Phase 1, but the reduction appeared smaller for the RI+ schedules. There was no effect of devaluation on these responses. A three-factor mixed-model ANOVA (group x phase x component) revealed a significant main effect of phase, $F(1,195)=167.95$, $p<.001$, MSe=33.42, $\eta^2_p=.463$[.194:.697], and component, $F(1,195)=25.56$, $p<.001$, MSe=15.60, $\eta^2_p=.116$[.006:.236], and a phase x group interaction, $F(1,195)=2.67$, $p=.034$, MSe=33.42, $\eta^2_p=.052$[.000:.165], but no other main effect nor interaction was significant, all *p*s> 0.060.

Group-mean within-bout rates are shown in the bottom panel of Figure 6. These data follow a similar pattern to those for the overall rates, but show a stronger numerical devaluation effect for all schedules, except the RI+Static schedule. A three-factor mixed-model ANOVA (group x phase x component) revealed that, except for component x group ($F<1$), all main effects and interactions were significant, all *p*s<.001, including the three-way interaction, $F(4,195)=5.03$, $p<.005$, MSe=1190.02, $\eta^2_p=.094$[.006:.197]. Simple effect analysis comparing the Phase 1 rates for the control and devalued components revealed no differences for any group, all *F*s<1. For phase 2 (extinction), there was a significantly lower rate for the devalued component for all schedules, smallest $F(1,195)=4.89$, $p=.035$, MSe=3959, $\eta^2_p=.102$[.002:.294], except for the RI+Static schedule, $F<1$.

These data corroborate and extend the suggestions made on the basis of the results of Experiment 4, and those findings reported by Garr et al. (2020). Devaluation effects on overall response rates are readily found for RR schedules, but not for RI schedules when the interval is long. This implies that responses reflect actions to the extent that there is some degree of variability on the rate of responding coupled with a corresponding variability in the rate of reinforcement. Thus, devaluation was pronounced for the RR and RI+Var schedules. However, on examination of the microstructure of responding, reinforcer devaluation was found to have no effect on the bout-initiation responding, as noted in the previous studies reported here. It did, however, impact the within-bout responding for those schedules where there was some degree of variability in responding. The one schedule on which there was little effect of devaluation was the RI+Static schedule. This effect is in line with the suggestions made by Garr et al. (2000), that the experienced nature of the response-reinforcer relationship is important in determining the degree to which actions are developed, as opposed to habits. It may also be worth comment that overall and within-burst response rates were higher when the feedback components (RI+) were added to the RI schedule. Although this did not make a great difference to the devaluation effects, overall, it is in line with the observations made by McDowell and Wixted (1988; see also Reed, 2015).

General Discussion

The current series of experiments were the first to explore the impact of reinforcer devaluation on human free-operant schedules with the aim of examining its effect on the microstructure of responding. Based on previous studies (e.g., Chen & Reed, 2023; 2024; Reed et al., 2018), it was predicted that reinforcer devaluation would differentially impact aspects of the microstructure of free-operant responding. It has been suggested that bout-initiation responses are controlled by the stimulus-value of the context (given by the reinforcement rate), and these bout-initiation responses may be similar to habits; whereas, within-bout responses are goal directed, and may be similar to actions (Falligant et al., 2024; Killeen, 2023; Reed, 2015; Shull, 2002). Given this, it was expected that reinforcer devaluation would have a greater effect on within-bout compared to bout-initiation responses. This account was corroborated in the current studies.

The current studies showed a variety of schedule effects in humans, including the differentiation of bout-initiation and within-bout responding by the reinforcer devaluation procedure (Dickinson et al., 1983). This also corroborates the effectiveness of such a devaluation manipulation for humans (see Allman et al., 2010). In terms of the overall rate of responding, it was clear that responding was greater to RR schedules than to RI schedules (Ferster & Skinner, 1957; Reed et al., 2018), even over a range of reinforcement rates (cf. Experiments 1a, 1b, 2, and 3). It was also shown that placing a positive response-reinforcer feedback loop onto an RI schedule increased overall response rates (McDowell & Wixted, 1986).

Importantly, it was shown that overall rates of reinforcement generally were more impacted after reinforcer devaluation on RR schedules than on RI schedules (Dickinson, 1985), although there were some limits to generality of this effect. The reinforcer devaluation effect was not exclusively shown on RR schedules, and was seen for some RI schedules. Devaluation occurred on both types of schedules when the microstructure of responding was examined – with within-bout responses being affected on most schedules, and bout-initiation responding not being affected on any schedule. These data suggest that within-bout responses, irrespective of the operative schedule, may be regarded as being akin to actions, and bout-initiation responses are more habitual in nature (Killeen, 2023; Reed et al., 2018). This is in accord with suggestions made based on the impact of manipulations aimed at enhancing the awareness of contingencies (Chen & Reed, 2023; 2024). Such a view suggests that responding is initiated due to the status of the context in which conditioning is occurring, and these responses are not necessarily goal directed but are elicited. Only when responding has commenced do the responses become goal directed. Perhaps this could be due to the first response in a bout activating a reinforcer representation, and allowing responding to become directed at that goal. This is similar to suggestions made about re-exposure to the reinforcer enhancing devaluation effects (Lopez, Balleine, & Dickinson, 1992).

It is not always the case that within-bout responding (i.e. that responding which occurs following the first response in a bout) is susceptible to devaluation effects. For example, on longer RI schedules (Experiments 1b and 3), shorter RR schedules (Experiment 1a), and RI+ schedules following training to show less variable responding (Experiment 5), within-bout responding is not shown to be sensitive to reinforcer devaluation. In some cases, RR schedules that do not support great experienced variation in the response-reinforcer relations, such as short RR that have little scope for variation, show smaller devaluation effects (Experiment 2). On the other hand, RI schedules that are short, and allow some degree of experience variation between response rate and reinforcement rate show devaluation (Experiment 3). When a positive feedback loop is placed onto an RI schedule, to create an RI+ schedule, then devaluation effects are noted (Experiment 4). When steps are taken to reduce this experienced variable, by pretraining static rates of responding, then devaluation does not have such a great impact (Experiment 5). The above pattern of results is in line with suggestions that variability in responding is important for the development of actions (Garr et al., 2020), and that this will occur on any schedule given some response-reinforcer relationship and a degree of variability in responding.

These suggestions regarding the susceptibility of bout-initiation and within-bout responding to devaluation are not necessarily at odds with previous views of the impact of schedules of reinforcement on the development of actions and habits (Dickinson, 1985; Perez & Dickinson, 2020). These previous views have suggested that contingencies where there is a strong response-reinforcer association, such as RR schedules, will tend to favour the development of goal-directed actions. In contrast, schedules where the response-reinforcer association is weaker, such as on RI schedules, will tend to allow the development of habits more readily (see also Garr et al., 2020). The current account differs in that it suggests that all schedules support both habits (bout-initiations) and actions (within-bout responding). The degree to which overall devaluation effects are noted will depend on the relative proportions of bout-initiations and within-bout responding.

Throughout the experiments reported in this current series, the cost of an investment response (i.e. one currency unit per investment) was maintained across both the initial training (Phase 1), and the after the devaluation in phase 2. This means that, relative to the exchange for GPB, a cost of an investment in the devalued currency is not as much as a cost of an investment in the non-devaluated currency in the other country. It is uncertain whether this would have an effect, and future studies may want to address this issue, although the manipulations to test this would make cross condition comparisons equally unbalanced in other ways.

In summary, the current experiments demonstrated that different aspects of human free-operant responding are differentially sensitive to reinforcer devaluation. Bout-initiation responses are not so affected, but within-bout responding, under some conditions, are decreased. This suggested that bout-initiation responses can be considered as habits (perhaps stimulus-driven), and within-bout responses can be considered as equivalent to actions and goal-directed. The conditions under which actions are supported also depend on the nature of the response-reinforcer contingency, and the variability of the responding that allows exposure to that correlation between response and reinforcement rate.

References

Allman, M.J., DeLeon, I.G., Cataldo, M. F., Holland, P. C., & Johnson, A.W. (2010). Learning processes affecting human decision making: An assessment of reinforcer selective Pavlovian-to-instrumental transfer following reinforcer devaluation. *Journal of Experimental Psychology: Animal Behavior Processes*, *36*(3), 402.

Baum, W.M. (1993). Performances on ratio and interval schedules of reinforcement: Data and theory. *Journal of the Experimental Analysis of Behavior*, *59*(2), 245-264.

Bouton, M.E. (2021). Context, attention, and the switch between habit and goal-direction in behavior. *Learning & Behavior*, *49*(4), 349-362.

Bradshaw, C.A., & Reed, P. (2012). Relationship between contingency awareness and human performance on random ratio and random interval schedules. *Learning and Motivation*, *43*(1-2), 55-65.

Chevée, M., Kim, C. J., Crow, N., Follman, E.G., Leonard, M.Z., & Calipari, E.S. (2023). Food restriction level and reinforcement schedule differentially influence behavior during acquisition and devaluation procedures in mice. *Eneuro*, *10*(9).

Chen, X., & Reed, P. (2023). The effect of brief mindfulness training on the micro-structure of human free-operant responding: Mindfulness affects stimulus-driven responding. *Journal of Behavior Therapy and Experimental Psychiatry*, *79*, 101821.

Chen, X., & Reed, P. (2024). Effect of instructions on the microstructure of human schedule performance. *Journal of Experimental Psychology: Animal Learning and Cognition*, *50*(1), 56.

Chen, X., Shui, X., Chen, J., Zhang, D., & Reed, P. (2025). Precursor processes of human schedule responding: Goal-directedness emerges after bout initiation. *Manuscript under review*.

Crossman, E.K., Bonhem, E.J., & Phelps. B.J. (1987). A comparison of response patterns on fixed-, variable-, and random-ratio schedules. *Journal of the Experimental Analysis of Behavior*, *48*, 395-406.

Dack, C., McHugh, L., & Reed, P. (2009). Generalization of causal efficacy judgments after evaluative learning. *Learning & Behavior*, *37*(4), 336-348.

DeRusso, A., Fan, D., Gupta, J., Shelest, O., Costa, R.M., & Yin, H.H. (2010). Instrumental uncertainty as a determinant of behavior under interval schedules of reinforcement. *Frontiers in Integrative Neuroscience*, *4*, 1737.

Dickinson, A. (1985). Actions and habits: the development of behavioural autonomy. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *308*(1135), 67-78.

Dickinson, A., Nicholas, D.J., & Adams, C.D. (1983). The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *The Quarterly Journal of Experimental Psychology*, *35*(1), 35-51.

Falligant, J.M., Hagopian, L.P., & Newland, M.C. (2024). Bouts, pauses, and units of operant performance: A primer. *Perspectives on Behavior Science*, *47*(3), 643-674.

Ferster, C.B., & Skinner, B.F. (1957) *Schedules of reinforcement*. New York, NY: Appleton.

Garr, E., & Delamater, A.R. (2019). Exploring the relationship between actions, habits, and automaticity in an action sequence task. *Learning & Memory*, *26*(4), 128-132.

Killeen, P.R. (2023). Theory of reinforcement schedules. *Journal of the Experimental Analysis of Behavior*, *120*(3), 289-319.

Kosaki, Y., & Dickinson, A. (2010). Choice and contingency in the development of behavioral autonomy during instrumental conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, *36*(3), 334.

Leander, J.D., Lippman, L.G., and Meyer, M.E. (1968). Fixed-interval performance as related to subjects' verbalizations of the reinforcement contingency. *Psychological Record*, 469-474.

Lopez, M., Balleine, B., & Dickinson, A. (1992). Incentive learning following reinforcer devaluation is not conditional upon the motivational state during re-exposure. *The Quarterly Journal of Experimental Psychology Section B*, *45*(4b), 265-284.

McDowell, J.J., & Wixted, J.T. (1986). Variable-ratio schedules as variable-interval schedules with linear feedback loops. *Journal of the Experimental Analysis of Behavior*, *46*(3), 315-329.

Mellgren, R.L., & Elsmore, T.F. (1991). Extinction of operant behavior: An analysis based on foraging considerations. *Animal Learning & Behavior*, *19*(4), 317-325.

Page, S., & Neuringer, A. (1985). Variability is an operant. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*(3), 429.

Pérez, O.D., Aitken, M.R., Zhukovsky, P., Soto, F.A., Urcelay, G.P., & Dickinson, A. (2019). Human instrumental performance in ratio and interval contingencies: A challenge for associative theory. *Quarterly Journal of Experimental Psychology*, *72*(2), 311-321.

Perez, O.D., & Dickinson, A. (2020). A theory of actions and habits: The interaction of rate correlation and contiguity systems in free-operant behavior. *Psychological Review*, *127*(6), 945.

Raia, C. P., Shillingford, S. W., Miller Jr, H. L., & Baier, P. S. (2000). Interaction of procedural factors in human performance on yoked schedules. *Journal of the Experimental Analysis of Behavior*, *74*(3), 265-281.

Randell, J., Ranjith-Kumar, A. C., Gupta, P., & Reed, P. (2009). Effect of schizotypy on responding maintained by free-operant schedules of reinforcement. *Personality and Individual Differences*, *47*(7), 783-788.

Reed, P. (1999). Effect of perceived cost on judgments regarding the efficacy of investment. *Journal of Economic Psychology*, *20*(6), 657-676.

Reed, P. (2007). Response rate and sensitivity to the molar feedback function relating response and reinforcement rate on VI+ schedules of reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes*, *33*(4), 428.

Reed, P. (2015). The structure of random ratio responding in humans. *Journal of Experimental Psychology: Animal Learning and Cognition*, *41*(4), 419.

Reed, P. (2023). Focused-attention mindfulness increases sensitivity to current schedules of reinforcement. *Journal of Experimental Psychology: Animal Learning and Cognition*, *49*(2), 127.

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*(3), 242-261.

Reed, P., Smale, D., Owens, D., & Freegard, G. (2018). Human performance on random interval schedules. *Journal of Experimental Psychology: Animal Learning and Cognition*, *44*(3), 309.

Shull, R.L. (2011). Bouts, changeovers, and units of operant behavior. *European Journal of Behavior Analysis*, *12*(1), 49-72.

Shull, R.L., Gaynor, S.T., & Grimes, J.A. (2002). Response rate viewed as engagement bouts: Effects of relative reinforcement and schedule type. *Journal of the Experimental Analysis of Behavior*, *75*(3), 247-274.

Sibley, R.M., Nott, H.M.R., & Fletcher, D.J. (1990). Splitting behaviour into bouts. *Animal Behaviour*, *39*, 63–69.

Thrailkill, E.A., & Bouton, M.E. (2015). Contextual control of instrumental actions and habits. *Journal of Experimental Psychology: Animal Learning and Cognition*, *41*(1), 69.

Thrailkill, E.A., Trask, S., Vidal, P., Alcalá, J.A., & Bouton, M.E. (2018). Stimulus control of actions and habits: A role for reinforcer predictability and attention in the development of habitual behavior. *Journal of Experimental Psychology: Animal Learning and Cognition*, *44*(4), 370.

Turner, K.M., & Balleine, B.W. (2024). Stimulus control of habits: Evidence for both stimulus specificity and devaluation insensitivity in a dual-response task. *Journal of the Experimental Analysis of Behavior*, *121*(1), 52-61.

Figure 1. Experiment 1a (RR-25). Response rates for devalued and control components. Top panel = overall response rates. Middle panel = bout-initiation rates. Bottom panel = within-but rates. Error bars = 95% confidence limits.

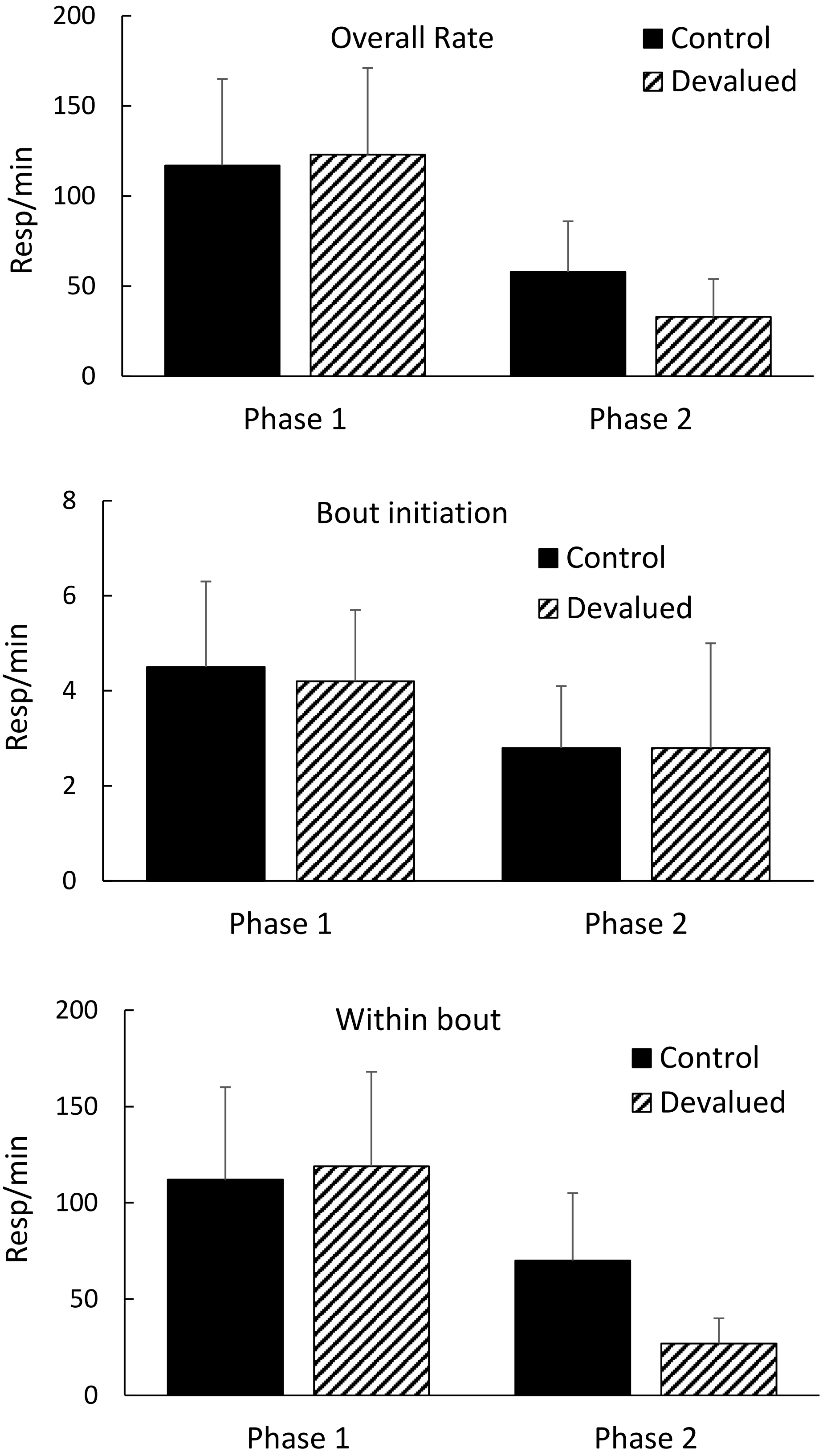

Figure 2. Experiment 1b (RI-15s). Response rates for devalued and control components. Top panel = overall response rates. Middle panel = bout-initiation rates. Bottom panel = within-bout rates. Error bars = 95% confidence limits.

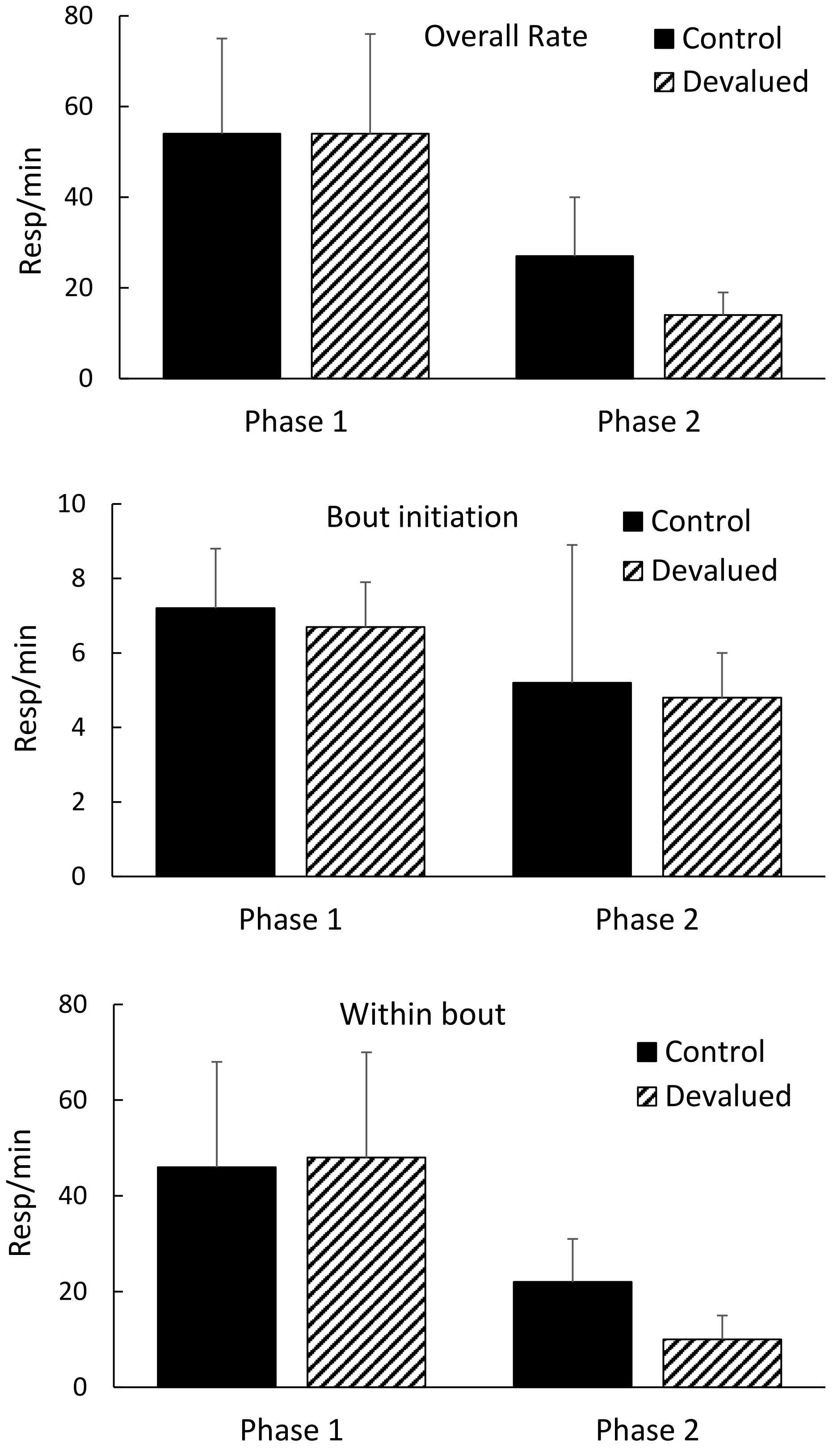

Figure 3. Experiment 2. Response rates for devalued and control countries responding on RR-10, RR-30, and RR-60 schedules. Top panel = overall response rates. Middle panel = bout-initiation rates. Bottom panel = within-bout rates. Error bars = 95% confidence limits.

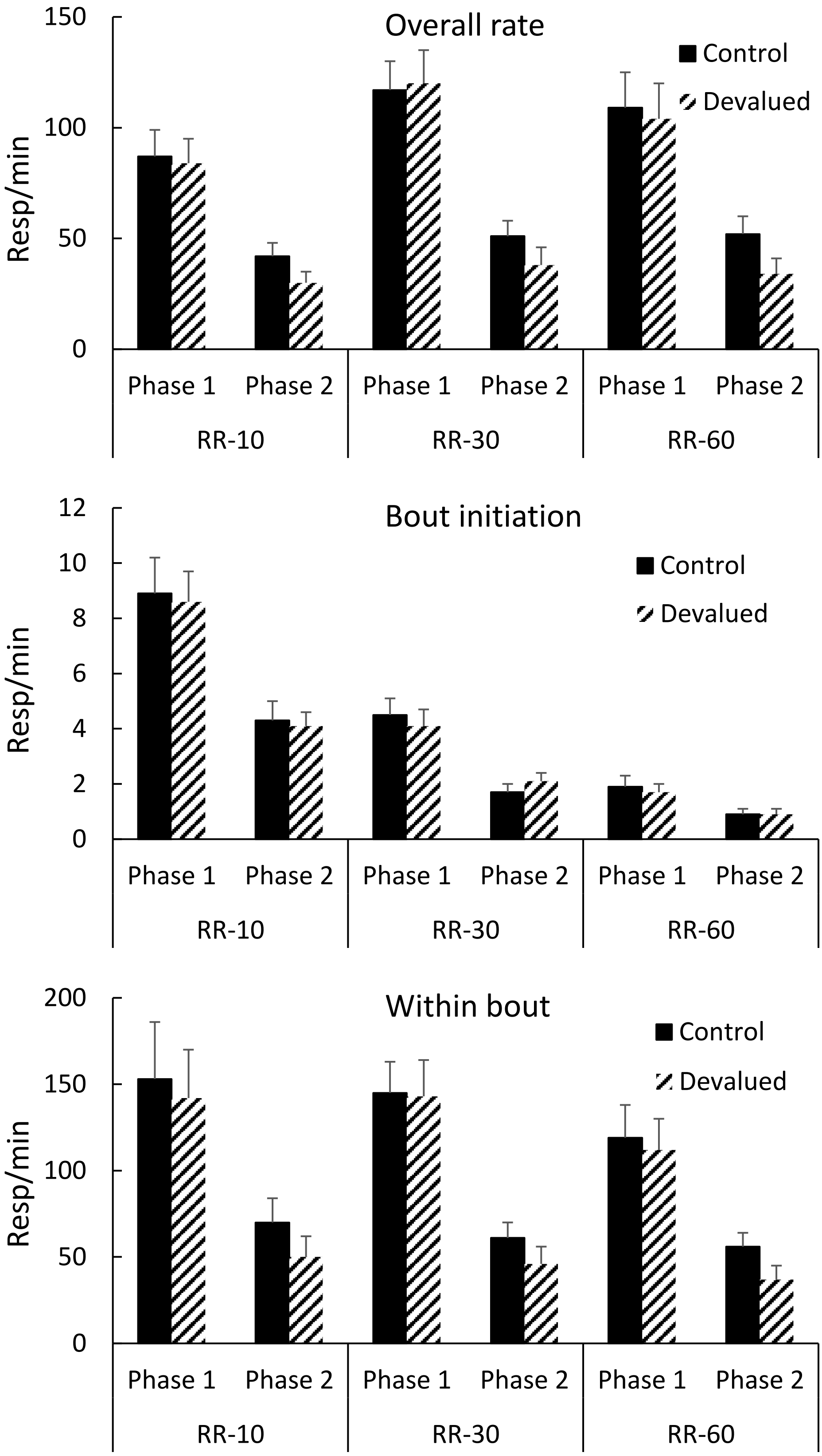

Figure 4. Experiment 3. Response rates for devalued and control countries responding on RI-10s, RI-30s, and RI-60s schedules. Top panel = overall response rates. Middle panel = bout-initiation rates. Bottom panel = within-bout rates. Error bars = 95% confidence limits.

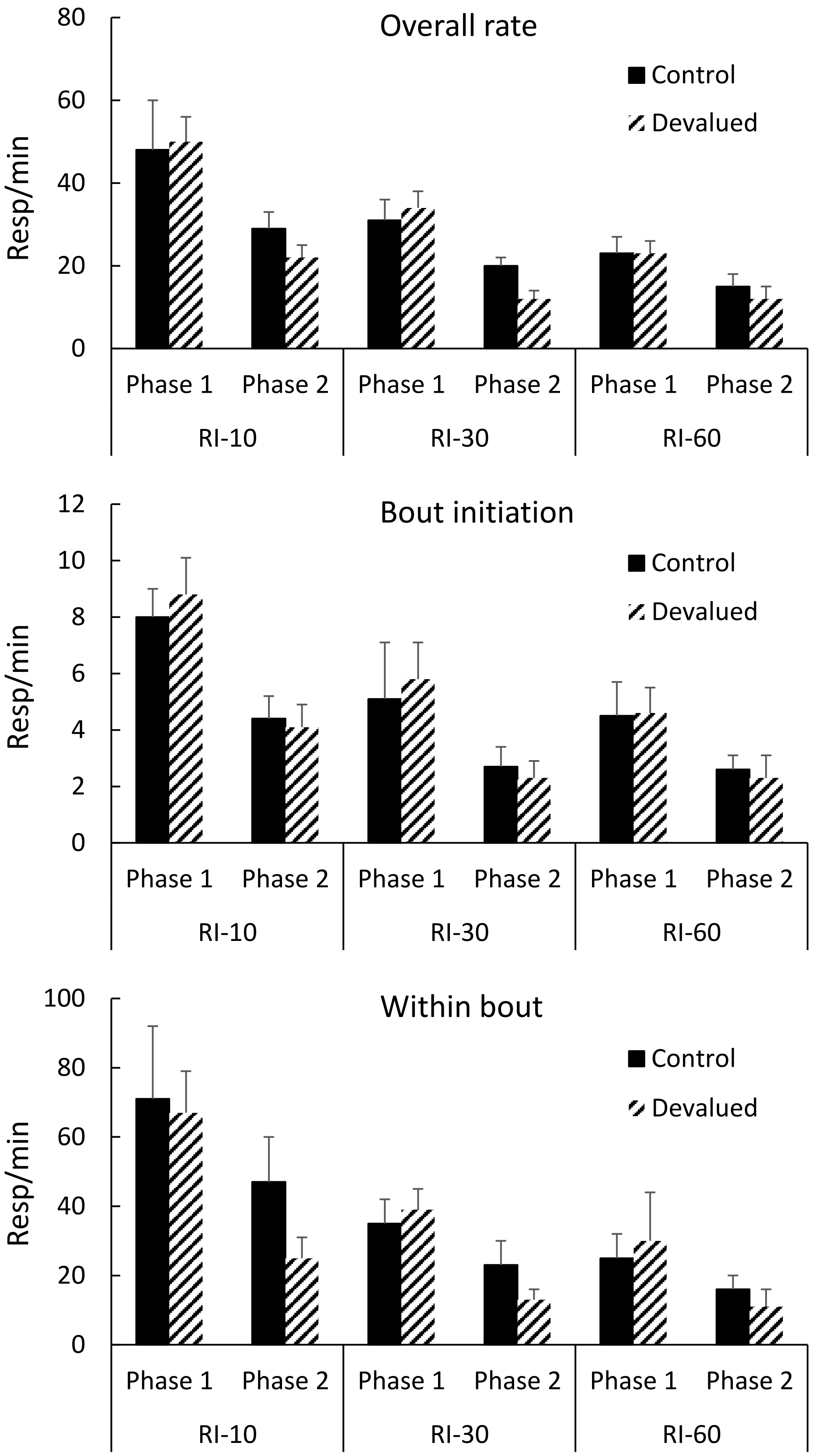

Figure 5. Experiment 4. Response rates for devalued and control components responding on RR, RI, and RI+ schedules. Top panel = overall response rates. Middle panel = bout-initiation rates. Bottom panel = within-bout rates. Error bars = 95% confidence limits.

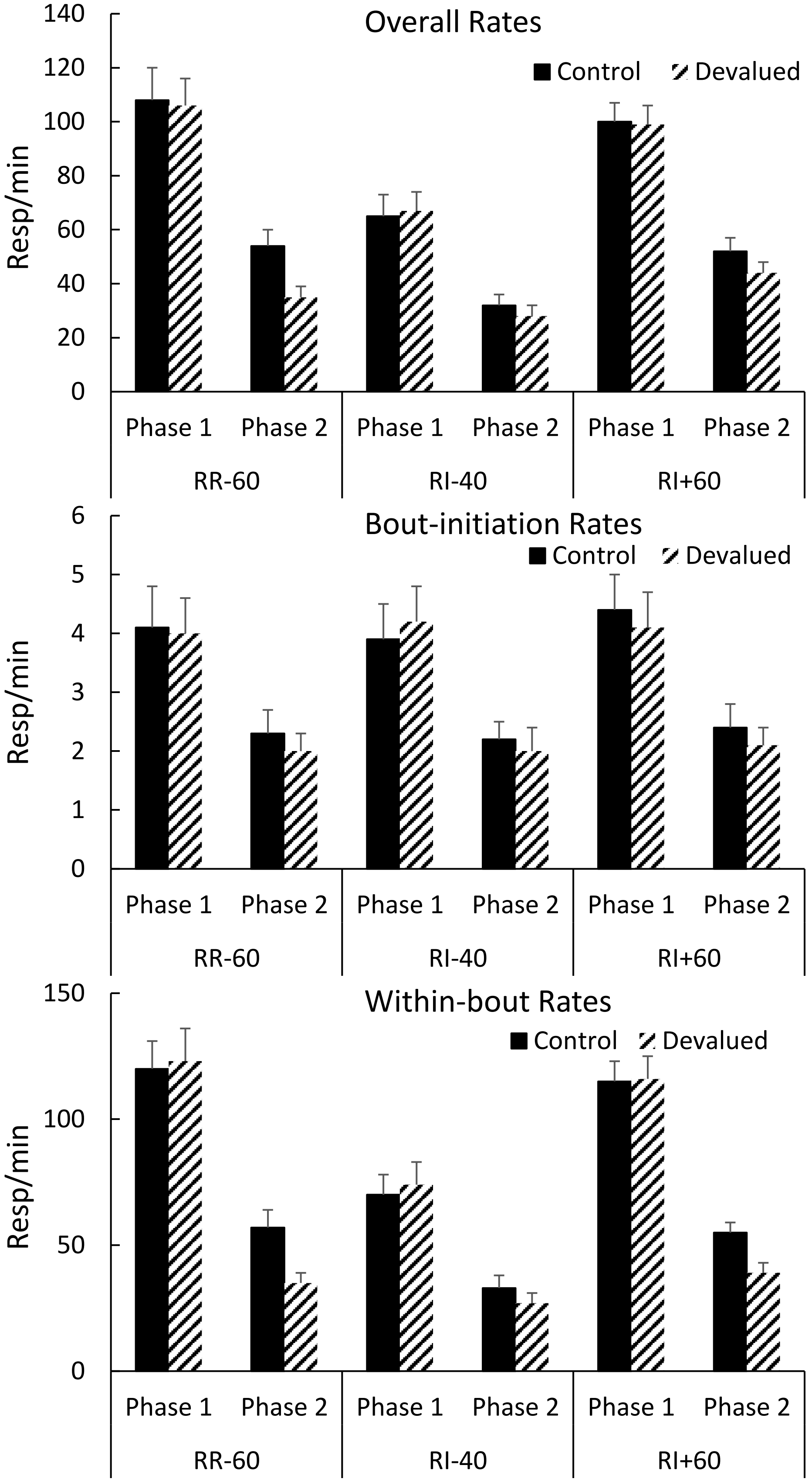

Figure 6. Experiment 5. Response rates for devalued and control components responding on RR, RI, RI+, RI+VAR, and RI+Static schedules. Top panel = overall response rates. Middle panel = bout-initiation rates. Bottom panel = within-bout rates. Error bars = 95% confidence limits.