

Goal-directed control in Pavlovian-instrumental transfer

Tina Seabrooke¹, Lee Hogarth², C. E. R. Edmunds¹ and Chris J. Mitchell¹

¹University of Plymouth

²University of Exeter

Running head: Goal-directed cue-control

Please address correspondence to:

Dr Tina Seabrooke

School of Psychology

University of Plymouth

Devon

PL4 8AA

United Kingdom

Email: tina.seabrooke@plymouth.ac.uk

Abstract

The current article concerns human outcome-selective Pavlovian-instrumental transfer (PIT), where Pavlovian cues selectively invigorate instrumental responses that predict common rewarding outcomes. Several recent experiments have observed PIT effects that were insensitive to outcome devaluation manipulations, which has been taken as evidence of an automatic “associative” mechanism. Other similar studies observed PIT effects that were sensitive to devaluation, which suggests a more controlled, goal-directed process. Studies supporting the automatic approach have been criticised for using a biased baseline, while studies supporting the goal-directed approach have been criticised for priming multiple outcomes at test. The current experiment addressed both of these issues. Participants first learned to perform two instrumental responses to earn two outcomes each (R1-O1/O3, R2-O2/O4), before four Pavlovian stimuli (S1-S4) were trained to predict each outcome. One outcome that was paired with each instrumental response (O3 and O4) was then devalued, so that baseline response choice at test would be balanced. Instrumental responding was then assessed in the presence of each individual Pavlovian stimulus, so that only one outcome was primed per trial. PIT effects were observed for the valued outcomes, $t_s > 3.99$, $p_s < .001$, but not for the devalued outcomes, $F < 1$, $BF_{10} = 0.29$. Hence, when baseline response choice was equated and only one outcome was primed per test trial, PIT was sensitive to outcome devaluation. The data therefore support goal-directed models of PIT.

Keywords: Pavlovian-instrumental transfer; outcome devaluation; goal-directed control

The ability to flexibly adjust learned behaviours based on current needs and desires is essential for adaptive decision-making. It is well established that both human and non-human animals readily exploit such *goal-directed* control (de Wit & Dickinson, 2009). Cues in our environment can also influence these instrumental behaviours, effectively directing responses towards predictable rewards in our surroundings (Cartoni, Balleine, & Baldassarre, 2016). In the laboratory, Pavlovian-instrumental transfer (PIT) tasks are often used to measure such cue-controlled reward-seeking (Colwill & Rescorla, 1988; Estes, 1943).

Traditional PIT tasks involve separate Pavlovian and instrumental training phases, followed by a transfer test. Two Pavlovian stimuli (S1 and S2) and two instrumental responses (R1 and R2) are first trained to predict different rewarding outcomes (O1 and O2) in separate Pavlovian (S1-O1, S2-O2) and instrumental (R1-O1, R2-O2) training phases. The Pavlovian cues are then presented and instrumental response choice is tested in extinction. A PIT effect can be seen when the Pavlovian cues selectively invigorate the response that shares a common outcome, relative to a baseline period. That is, stimulus S1 usually increases response R1 (both paired with O1), and S2 increases R2 (both paired with O2). Thus, outcome-selective PIT procedures demonstrate that reward-predictive cues can motivate instrumental responses that predict common outcomes (Colwill & Rescorla, 1988).

Several experiments with both rats and humans have found that PIT effects are robust against outcome devaluation manipulations, which is regarded as evidence of an automatic, non-goal-directed mechanism (e.g., Hogarth & Chase, 2011; Rescorla, 1994; Watson, Wiers, Hommel, & de Wit, 2014). Such studies usually employ a traditional PIT procedure (as described above), but devalue one of the outcomes (e.g., O1) immediately before the test. On test, overall responding is usually biased towards the response that produced the still-valued outcome (i.e., R2), which demonstrates that *overall* response choice is goal-directed (de Wit and Dickinson, 2009). In addition, the Pavlovian cues typically increase instrumental choice

of the response that was paired with the same outcome, regardless of whether that outcome is valued or not. Most importantly, the PIT effects for the valued and devalued outcomes do not usually significantly differ in size. Thus, although overall response choice is sensitive to devaluation, the effect of Pavlovian cues on instrumental response choice is often *insensitive* to changes in outcome value (for a review, see Watson et al., 2018).

The insensitivity of PIT to devaluation described above is consistent with link-based, “associative” theories of the phenomenon, in which the stimulus is suggested to automatically trigger the instrumental response that shares a common outcome, without consideration of the value of that outcome (e.g., Cohen-Hatton et al., 2013; Watson et al., 2014). Such insensitivity appears to be especially *inconsistent* with goal-directed models of PIT, which instead suggest that PIT effects are the product of controlled, higher-order cognition (Hardy, Mitchell, Seabrooke, & Hogarth, 2017; Seabrooke, Le Pelley, Porter, & Mitchell, 2018). According to this goal-directed account of PIT, participants infer that the presentation of a Pavlovian cue during the transfer test signals an increased availability of the associated outcome. This inference then leads the participants to preferentially perform the instrumental response that was paired with the cued outcome in a *non-automatic* fashion (Hogarth et al., 2014; Seabrooke, Hogarth, & Mitchell, 2016). While this goal-directed account has intuitive appeal, critics argue that the observed insensitivity to devaluation undermines it; why would participants deliberately choose a devalued outcome, whatever its perceived availability?

We propose that there are intrinsic methodological issues with the standard PIT design that render the usual interpretation of those studies premature. In most PIT experiments, the effect of the Pavlovian cues on instrumental responding is assessed relative to baseline periods in which no Pavlovian cues are present (e.g., Hogarth & Chase, 2011; Watson et al., 2014). As noted above, instrumental responding is generally biased towards the response that signals the still-valued outcome, which is good evidence that the devaluation

manipulation was effective. However, a biased baseline has important implications for the interpretation of PIT effects. If participants tend to choose the *valued* outcome at baseline, then the opportunity to observe a PIT effect for that valued outcome is reduced, due to a ceiling effect. Conversely, minimal responding for the *devalued* outcome at baseline allows more opportunity to observe a PIT effect for that devalued outcome, because baseline responding is closer to floor. When PIT effects for valued and devalued outcomes do not significantly differ in size, then, it might simply be because there is more scope to see a PIT effect for the devalued outcome than the valued outcome.

We recently attempted to assess the effect of outcome devaluation on PIT when baseline response choice was balanced (Seabrooke, Le Pelley, Hogarth, & Mitchell, 2017). In our second experiment, for example, participants first learned to perform two instrumental responses to earn points towards two different food outcomes each (R1-O1/O3; R2-O2/O4). One outcome that was paired with each response (O3 and O4) was then devalued by coating it in a ground clove paste – a procedure that rendered those outcomes very unpleasant. Notably, the design balanced baseline response choice after this devaluation manipulation, because each response was now paired with one valued (O1/O2) and one devalued (O3/O4) outcome. Instrumental response choice was then tested in the presence of stimulus compounds that contained pictorial stimuli (S1-S4) that were associated with the outcomes. S1 and S4 were presented on some trials; S2 and S3 were presented on the remainder. Importantly, these stimulus compounds both signalled one outcome that was associated with each instrumental response, as well as one valued and one devalued outcome. The S1+S4 compound, for example, signalled the valued O1 (paired with R1) and the devalued O4 (paired with R2). Insensitivity to devaluation would have been revealed if these stimulus compounds did not bias response choice in either direction, because both responses were primed equally. Goal-directed control, by contrast, would have been revealed by a selective

bias towards the cued, valued outcome (e.g., respond R1 rather than R2 in the presence of S1+S4). The results provided strong evidence of goal-directed control; each stimulus compound biased instrumental response choice towards the cued and valued outcome.

Seabrooke et al.'s (2017) results are exactly what would be predicted by goal-directed models of PIT. The results have, however, been disputed by Watson et al. (2018), who questioned the relevance of Seabrooke et al.'s findings to the PIT phenomenon more generally. Watson et al. suggested that presenting two stimuli together on test (e.g., S1+S4) - and hence priming both valued and devalued outcomes simultaneously - might have encouraged a goal-directed process that is not responsible for other, more typical PIT effects. They also suggested that task instructions might have encouraged a controlled process in previous experiments that observed sensitivity to devaluation (Allman, DeLeon, Cataldo, Holland, & Johnson, 2010; Eder & Dignath, 2016a, 2016b). In this way, the authors suggested that the previous demonstrations of sensitivity to devaluation might not be indicative of the mechanisms underlying the PIT phenomenon broadly speaking, but are quite specific to those procedures. The present experiment tested this possibility.

The current study retained the key components of the traditional PIT procedure, but adapted the instrumental training so that baseline responding would be balanced at test. Table 1 shows the design. Following Seabrooke et al. (2017), participants first learned to perform two instrumental responses (R1 and R2) to earn points towards two different food outcomes each: crisps, popcorn, cashew nuts and nachos (R1-O1/O3, R2-O2/O4). Four Pavlovian stimuli (S1-S4) were then trained to predict each outcome. Next, one outcome that was paired with each response (O3 and O4) was devalued, so that baseline response choice would remain balanced. On test, the Pavlovian cues were presented individually and participants were able to freely perform both instrumental responses.

We expected the cues that signalled the valued outcomes (S1 and S2) to selectively increase the instrumental response (R1/R2) that shared a common outcome (O1/O2). The more interesting result concerns the devalued outcomes. If PIT effects are only sensitive to devaluation manipulations when multiple outcomes are cued, as proposed by Watson et al. (2018), then we should now see *insensitivity* to devaluation. That is, S3 and S4 should increase R1 and R2, respectively, and the PIT effects for the valued and devalued outcomes should not significantly differ in size. In contrast, if PIT is goal-directed and previous demonstrations of insensitivity to devaluation were the consequence of a biased baseline, then we should now see a PIT effect for the valued outcomes, but not the devalued outcomes. Hence, the current experiment addresses the problem with standard PIT procedures in that there should be no baseline bias at test. At the same time, it addresses Watson et al.'s concern with Seabrooke et al.'s (2017) study, in that only one outcome is cued per test trial. The experiment therefore provides a crucial evaluation of the two classes of theory that seek to explain human PIT: the “associative” link account and the goal-directed cognition account.

Method

Participants. Forty University of Plymouth psychology undergraduates (33 females, aged between 18 and 39, $M = 20.18$ years, $SEM = 0.50$ years) took part for course credit. The participants were screened for food allergies and intolerances. The University of Plymouth Ethics Committee approved the experiment.

Apparatus and materials. The experiment was programmed in E-Prime 2.0 and was presented on a 22-inch (55.88cm) computer monitor. The text was presented in white on a black screen, and responses were made using a standard keyboard and mouse. The participants wore headphones throughout the experiment. Unopened bags of Tyrell's lightly sea salted crisps (150g), Tyrell's sea salted popcorn (70g), Sainsbury's salted jumbo cashew nuts (400g) and Doritos' original nachos (180g) served as outcome props. These brands were

also used for the devaluation manipulation; the foods were decanted into separate transparent containers, with the food name written on the lid. For the devaluation manipulation, ground cloves were combined with olive oil (11g oil per 5g cloves) to form a paste that was brushed heavily onto the devalued foods. The valued outcomes were simply transferred from their original packaging to their containers. The four foods were randomly assigned to the roles of outcomes O1-O4 for each participant.

Procedure. Before the experiment, the participants were warned that they would be asked to sample foods during the experiment, that they might not match their expectations, and that they might taste unpleasant. The food props were presented, and participants were told they could win points toward them during the experiment. The participants gave liking ratings (i.e., rated how much they wanted to eat each food) for each outcome in a random order (1 = “Not at all”, 7 = “Very much”), before the food props were removed.

Instrumental training. Before the instrumental training phase, the participants were told that they could earn the four foods by pressing the “E” and “I” keys on the keyboard, and that they might need to press the keys several times to earn rewards. They were told that each button earned two foods each, and that their task was to learn which keys earned each food. Finally, they were told that they did not need to press the keys in a particular order, but that they did need to learn about both keys, so they should try to press them roughly equally.

Participants were allowed to make as many R1 and R2 responses as they wished in a continuous fashion. Each key press (R1 or R2) had a 0.2 probability of being reinforced. This was programmed as a series of “trials” in the following way. Each trial began with the presentation of a choice symbol (“← or →”), which remained until the participant selected the E or I key. The two keys were counterbalanced with respect to their roles as R1 and R2. The screen changed to a blank screen (100ms) after each response, and each response had a

20% chance of producing an outcome. When an outcome was not available, the program looped back to the start of the trial (and hence participants had to perform another instrumental response) until an outcome *was* available. Outcomes were presented as points rather than real food rewards to avoid a generalised devaluation of the outcomes through satiation (see Colagiuri & Lovibond, 2015). Outcomes were presented for 1500ms, with the text “You earn one [REWARD] point” placed above a picture of the corresponding food outcome. The text in brackets was replaced with the appropriate outcome (e.g., “CRISPS”).

There were 32 instrumental training trials. Pilot testing revealed that many participants failed to learn the contingencies with a free responding procedure. We therefore adopted a staggered method. For the first eight trials, only one outcome that was associated with each instrumental response was available. For half of the participants, R1 produced O1 and R2 produced O2 (when an outcome was available). The *other* outcomes were available for the next eight trials (R1-O3, R2-O4). The contingencies were revealed in the opposite order for the remaining participants (i.e., R1-O3, R2-O4 first, then R1-O1, R2-O2). This staggered procedure allowed us to establish the instrumental contingencies gradually, which we hoped would aid learning. On the remaining 16 trials, two outcomes (one associated with each instrumental response) were available on each trial, and training of the four instrumental contingencies (see Table 1) was intermixed. Hence, the R1-O1 and R2-O2 relationships were trained on a random eight trials, and the R1-O3 and R2-O4 relationships were trained on the rest. The trials were separated by 500ms intervals.

Pavlovian training. Before the Pavlovian training phase, the participants were told that different colours would predict points towards the different food rewards, and that their task was to learn which colour predicted each reward. Each trial began with the central presentation of a Pavlovian cue. Red, blue, green and orange rectangles were randomly assigned to the roles of stimulus S1-S4 for each participant. The question, “Which reward

will follow?” was superimposed on the Pavlovian stimulus, above the four outcome options (“CRISPS”, “POPCORN”, “CASHEWS” and “NACHOS”). The four options were arranged vertically and were ordered randomly on each trial, and participants selected an option using the mouse (see e.g., Watson, Wiers, Hommel, Gerdes, & de Wit, 2017, for a similar procedure). The correct outcome (see Table 1) was then presented, regardless of the participant’s response. Outcomes were superimposed on the Pavlovian cue for 1500ms, with the text “[Stimulus colour] earns one [REWARD] point” placed above a picture of the corresponding food outcome. The text in brackets was replaced by the appropriate stimulus colour (e.g., “Blue”) and reward (e.g., “CRISPS”). Incorrect predictions were followed by error noises. There were 12 blocks of four trials (48 trials total), with each Pavlovian cue presented once per block in a random order. The trials were separated by 500ms intervals. Pavlovian contingency knowledge was assessed immediately after the Pavlovian training phase; the four Pavlovian cues (S1-S4) were presented in a random order, and the participants had to select the outcome that the cue predicted.

Booster instrumental training. Thirty-two booster instrumental training trials were administered to reinforce the instrumental contingencies. They followed the procedure of the first instrumental training session, except that the four contingencies were intermixed from the start. Thus, two instrumental contingencies (e.g., R1-O1, R2-O2) were trained on a random half of the trials, and the other two contingencies (R1-O3, R2-O4) were trained on the rest. Afterwards, participants answered four instrumental knowledge questions relating to which response earned points towards each outcome, and gave confidence ratings for each (1 = “Not at all confident”, 7 = “Very confident”).

Outcome devaluation. For the outcome devaluation procedure, participants were asked (but not forced) to sample the four food outcomes, and were told that those were the foods that were now available for them to win. The valued outcomes were always revealed

and sampled before the devalued outcomes, but were otherwise randomly sampled. Post-devaluation liking ratings were then taken as at the start of the experiment.

Transfer test. During the transfer test, a Pavlovian cue (S1, S2, S3 or S4) was presented centrally for six seconds on each trial, with the instrumental choice symbol superimposed on top. The participants were free to perform R1 and R2 as often as they liked. The choice symbol (“← or →”) disappeared for 100ms after each instrumental response. The stimulus trials were separated by inter-trial intervals (ITI), which were the same as the stimulus trials, except that no Pavlovian cues were presented. There were 32 trials, and every four trials included one presentation of each Pavlovian cue, presented in a random order. Rewards were not presented on test. The transfer test was preceded by a practice phase, which included one presentation of each of the four Pavlovian cues (in a random order). As in our previous work, the participants did not actually consume the food rewards at the end of the experiment for ethical reasons (Seabrooke et al., 2017). Finally, the participants completed another set of instrumental and Pavlovian knowledge tests after the transfer test.

Results

The trial-level raw data are publicly archived at <https://osf.io/9ptqv/>.

Every participant reported perfect contingency knowledge on the first instrumental test. On the final test, 65% of participants reported perfect knowledge of the instrumental contingencies. Average accuracy on the final block of Pavlovian training was 98.13% ($SEM = 1.38\%$). Collapsed across blocks, there was no significant effect of stimulus on accuracy, $F(3, 117) = 1.55, p = .21, \eta_g^2 = .02$. On the first and second Pavlovian knowledge tests, 92.50% and 95% of participants reported perfected Pavlovian contingency knowledge, respectively.

Figure 1a shows the mean liking ratings for the valued (O1/O2) and devalued (O3/O4) outcomes. Liking for the valued and devalued outcomes did not appear to differ

before devaluation, and the non-devalued outcomes were preferred after the devaluation phase. This general pattern was confirmed in a 2 (test: pre-devaluation vs. post-devaluation) \times 2 (outcome: valued vs. devalued) ANOVA, in which the crucial interaction was observed, $F(1, 39) = 258.57, p < .001, \eta_g^2 = .57$. Pre-devaluation liking ratings did not significantly differ for the valued and devalued outcomes, $t(39) = 1.26, p = .21, d_z = 0.20$. The valued outcomes received higher liking ratings than the devalued outcomes after devaluation, $t(39) = 28.63, p < .001, d_z = 4.53$.

Figure 1b shows the mean number of R1 and R2 responses during the transfer test. We first submitted the data to a 2 (cued response: R1 [S1/S3] vs. R2 [S2/S4]) \times 2 (cued outcome value: valued [S1/S2] vs. devalued [S3/S4]) \times 2 (instrumental response: R1 vs. R2) repeated-measures ANOVA. This initial analysis focuses on the cued trials and ignores the ITI trials. A significant main effect of cued outcome value was observed, with participants performing more responses overall in the presence of the stimuli that signalled the valued outcomes ($M = 7.56, SEM = 0.59$) than the stimuli that signalled the devalued outcomes ($M = 6.50, SEM = 0.49$), $F(1, 39) = 5.56, p = .02, \eta_g^2 = .007$. There was no significant main effect of cued response, $F(1, 39) = 1.26, p = .27, \eta_g^2 = .0001$, suggesting that the overall number of instrumental responses performed did not significantly differ in the presence of the stimuli that signalled R1 (S1/S3) and R2 (S2/S4). There was also no significant main effect of instrumental response, $F < 1$, demonstrating that the overall number of R1 and R2 responses did not significantly differ. There was a significant interaction between cued response and instrumental response, $F(1, 39) = 14.73, p < .001, \eta_g^2 = .08$, which suggests that the cues tended to selectively promote the instrumental response that was associated with a common outcome. However, neither the cued response \times outcome value interaction, $F(1, 39) = 3.52, p = .07, \eta_g^2 = .0008$, nor the cued outcome value \times instrumental response interaction, $F < 1$,

reached significance. Finally, and most importantly, there was a significant three-way interaction, $F(1, 39) = 15.78, p < .001, \eta_g^2 = .05$.

The significant three-way interaction prompted separate analyses comparing the effect of cued response and instrumental response for the stimuli that signalled the valued outcomes (S1/S2) and then for the stimuli that signalled the devalued outcomes (S3/S4). For the stimuli that signalled the valued outcomes, there was a trend towards participants performing more responses in the presence of S1 ($M = 7.82, SEM = 0.84$) than S2 ($M = 7.29, SEM = 0.83$), $F(1, 39) = 3.93, p = .05, \eta_g^2 = .002$. There was no significant main effect of instrumental response, $F < 1$. Most importantly, there was a significant cued response \times instrumental response interaction, $F(1, 39) = 25.04, p < .001, \eta_g^2 = .21$. Stimulus S1 produced more R1 than R2 responses, $t(39) = 3.96, p < .001, d_z = 0.63$. Stimulus S2, by contrast, elicited more R2 responses than R1 responses, $t(39) = 3.99, p < .001, d_z = 0.63$. Thus, outcome-selective PIT effects were observed for both stimuli (S1 and S2) that signalled valued outcomes.

A comparable analysis was conducted for the stimuli that signalled the devalued outcomes. This analysis revealed no significant main effects or interactions, $F_s < 1$. The non-significant interaction was supported by a Bayes Factor (BF_{10}) of 0.29 (JASP Team, 2018). Thus, PIT effects were not observed for the devalued outcomes.

We also examined instrumental responding during the stimulus periods relative to the ITI periods. The first point to note is that the number of R1 and R2 responses during the ITI periods did not significantly differ, $t < 1, BF_{10} = 0.19$. We also explored whether the cues elevated the overall level of responding (collapsed across R1 and R2) compared with the ITI periods. The stimuli that signalled the valued outcomes (S1/S2) did not significantly elevate overall levels of responding compared with the ITI, $t < 1, BF_{10} = 0.23$. The stimuli that

signalled the devalued outcomes (S3/S4), by contrast, significantly depressed overall levels of responding compared with the ITI, $t(39) = 2.94$, $p = .005$, $d_z = 0.46$.

Discussion

When baseline response choice was equated in an otherwise traditional PIT design, an outcome-selective PIT effect was observed for the valued outcomes, but not the devalued outcomes. Moreover, the stimuli that signalled the devalued outcomes produced a generalised suppression of instrumental responding, compared with the ITI periods. Together, these data suggest that PIT effects are sensitive to devaluation manipulations when baseline response choice is balanced. Our results complement previous demonstrations of PIT effects that were sensitive to devaluation (Allman et al., 2010; Eder & Dignath, 2016a, 2016b), and they suggest that past observations of *insensitivity* to devaluation were quite possibly the result of a biased baseline (e.g., Hogarth, 2012; Hogarth & Chase, 2011; Watson et al., 2014). Finally, the results suggest that Seabrooke et al.'s (2017) observed sensitivity to devaluation does not depend on multiple outcomes being cued together on test.

It might be argued that the current experiment artificially promoted controlled cognition by pairing each instrumental response during training with two outcomes rather than one (as in standard PIT experiments). We see no obvious a priori reason why learning two contingencies on separate training trials should promote controlled cognition at test. Furthermore, in light of the current results, insensitivity to outcome devaluation only seems to be observed under the precise and highly constrained conditions present in the original PIT procedure, where each response is paired with one outcome, and those outcomes are then primed individually on test. As argued above, the bias in baseline responding seen under

these conditions might conceal the effect of outcome value on PIT¹. Furthermore, we would argue that the current design, as well as previous experiments that primed multiple outcomes together on test (Seabrooke et al., 2017), is a better reflection of the Pavlovian-instrumental interactions that might occur in the complex environment that exists outside of the laboratory.

Our results have important theoretical implications, particularly for link-based accounts that specifically predict that PIT effects should be insensitive to changes in outcome value. S-O-R theory, for example, suggests that the presentation of a Pavlovian stimulus S during the transfer test activates the sensory properties (not the value) of the associated outcome O, which then automatically triggers the associated instrumental response (Watson et al., 2018). Mediated S-R theory similarly predicts that PIT effects should be immune to post-training changes in outcome value, because it proposes that PIT effects are mediated by a direct stimulus-response link that does not incorporate the outcome representation at all (Cohen-Hatton et al., 2013). The results do not support such theories. Rather, following Seabrooke et al. (2016, 2017), we propose that participants' actions during the transfer test are mediated by outcome value and perceived outcome availability. When a Pavlovian cue that signals a valued outcome is presented, we suggest that participants deliberately respond for that outcome because it is perceived to be highly available. When a stimulus that signals a devalued outcome is presented, by contrast, the participants have no real basis for a strong preference of one response over the other; they can either respond for the devalued outcome that is perceived to be readily available, or for a high-value outcome that is assumed to be unavailable (because it is not cued). We suspect that this dilemma underlies the chance performance that was observed when the devalued outcomes were cued.

¹ There are, of course, other ways to balance baseline response choice apart from pairing each response with an equal number of valued and devalued outcomes. For example, extinguishing the context should, in principle, balance baseline response choice in a similar way to as in the current experiment.

It is worth noting that Rescorla (1994) observed a PIT effect in rats that was seemingly insensitive to outcome devaluation, even though baseline response choice was balanced. This result opens the possibility that rodent and human PIT effects might be mediated by fundamentally different mechanisms. Alternatively, Rescorla's result could be due to an insufficiently strong devaluation manipulation (the devaluation check was that the animals left "some" of the devalued outcomes). It is also possible that, given the particularly complex training procedure that Rescorla used, the rats learnt that the Pavlovian stimuli and instrumental responses predicted outcomes, but not the specific relationships between each stimulus, response and outcome. This could have produced to the apparent insensitivity to devaluation that was observed. These possibilities warrant further attention.

Finally, the data are interesting in light of a recent experiment by Watson et al. (2017), who found that obese participants showed stronger PIT effects for high-calorie foods than low-calorie foods. The authors suggested that, because outcome value was established *before* training, obese participants might have formed weaker Pavlovian and instrumental links with the low-calorie foods than the high-calorie foods. From an associative link (S-O-R) perspective, weak Pavlovian cues should produce smaller PIT effects than strong Pavlovian cues. The current data, however, suggest that PIT effects are sensitive to outcome value even when outcome value is established *after* training (and hence equally strong associations should form between valued and devalued outcomes). We therefore suggest that Watson et al.'s (2017) findings might instead reflect a goal-directed decision-making process, where obese participants showed smaller PIT effects for low-calorie foods because those foods were less desired than the high-calorie foods on test.

To conclude, the current work builds on a rapidly expanding body of literature that explores the effect of outcome devaluation on PIT in human subjects. When baseline response choice was equated in an otherwise traditional PIT design, cue-elicited response

choice was highly sensitive to a strong devaluation manipulation. The results support goal-directed theories of PIT over stimulus-driven, automatic models. We also suggest that the insensitivity to devaluation that has been previously observed with the traditional PIT design might have arisen because of imbalances in baseline response choice.

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*, 402–408.
<https://doi.org/10.1037/a0017876>
- Baguley, T. (2012). Calculating and graphing within-subject confidence intervals for ANOVA. *Behavior Research Methods*, *44*, 158–175. <https://doi.org/10.3758/s13428-011-0123-7>
- Cartoni, E., Balleine, B., & Baldassarre, G. (2016). Appetitive Pavlovian-instrumental transfer: A review. *Neuroscience and Biobehavioral Reviews*, *71*, 829–848.
<https://doi.org/10.1016/j.neubiorev.2016.09.020>
- Cohen-Hatton, S. R., Haddon, J. E., George, D. N., & Honey, R. C. (2013). Pavlovian-to-instrumental transfer: Paradoxical effects of the Pavlovian relationship explained. *Journal of Experimental Psychology: Animal Behavior Processes*, *39*, 14–23.
<https://doi.org/10.1037/a0030594>
- Colagiuri, B., & Lovibond, P. F. (2015). How food cues can enhance and inhibit motivation to obtain and consume food. *Appetite*, *84*, 79–87.
<https://doi.org/10.1016/j.appet.2014.09.023>

- Colwill, R. M., & Rescorla, R. A. (1988). Associations between the discriminative stimulus and the reinforcer in instrumental learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *14*, 155–164. <https://doi.org/10.1037/0097-7403.14.2.155>
- de Wit, S., & Dickinson, A. (2009). Associative theories of goal-directed behaviour: a case for animal – human translational models. *Psychological Research*, *73*, 463–476. <https://doi.org/10.1007/s00426-009-0230-6>
- Eder, A. B., & Dignath, D. (2016a). Asymmetrical effects of posttraining outcome reevaluation on outcome-selective Pavlovian-to-instrumental transfer of control in human adults. *Learning and Motivation*, *54*, 12–21. <https://doi.org/10.1017/CBO9781107415324.004>
- Eder, A. B., & Dignath, D. (2016b). Cue-elicited food seeking is eliminated with aversive outcomes following outcome devaluation. *Quarterly Journal of Experimental Psychology*, *69*, 574–588. <https://doi.org/10.1080/17470218.2015.1062527>
- Estes, W. K. (1943). Discriminative conditioning. I. A discriminative property of conditioned anticipation. *Journal of Experimental Psychology*, *32*, 150–155. <https://doi.org/http://dx.doi.org/10.1037/h0058316>
- Hardy, L., Mitchell, C. J., Seabrooke, T., & Hogarth, L. (2017). Drug cue reactivity involves hierarchical instrumental learning: Evidence from a biconditional Pavlovian to instrumental transfer task. *Psychopharmacology*, *234*, 1977–1984. <https://doi.org/10.1007/s0021>
- Hogarth, L. (2012). Goal-directed and transfer-cue-elicited drug-seeking are dissociated by pharmacotherapy: Evidence for independent additive controllers. *Journal of Experimental Psychology: Animal Behavior Processes*, *38*, 266–278. <https://doi.org/10.1037/a0028914>

- Hogarth, L., & Chase, H. W. (2011). Parallel goal-directed and habitual control of human drug-seeking: Implications for dependence vulnerability. *Journal of Experimental Psychology: Animal Behavior Processes*, *37*, 261–276.
<https://doi.org/10.1037/a0022913>
- Hogarth, L., Retzler, C., Munafò, M. R., Tran, D. M. D., Troisi, J. R., Rose, A. K., ... Field, M. (2014). Extinction of cue-evoked drug-seeking relies on degrading hierarchical instrumental expectancies. *Behaviour Research and Therapy*, *59*, 61–70.
<https://doi.org/10.1016/j.brat.2014.06.001>
- JASP Team. (2018). JASP (Version 0.8.6)[Computer software]. Retrieved from <https://jasp-stats.org/>
- Rescorla, R. A. (1994). Transfer of instrumental control mediated by a devalued outcome. *Animal Learning & Behavior*, *22*, 27–33. <https://doi.org/10.3758/BF03199953>
- Seabrooke, T., Hogarth, L., & Mitchell, C. J. (2016). The propositional basis of cue-controlled reward seeking. *Quarterly Journal of Experimental Psychology*, *69*, 2452–2470. <https://doi.org/10.1080/17470218.2015.1115885>
- Seabrooke, T., Le Pelley, M. E., Hogarth, L., & Mitchell, C. J. (2017). Evidence of a goal-directed process in human Pavlovian-instrumental transfer. *Journal of Experimental Psychology: Animal Learning and Cognition*, *43*, 377–387.
<https://doi.org/10.1037/xan0000147>
- Seabrooke, T., Le Pelley, M. E., Porter, A., & Mitchell, C. J. (2018). Extinguishing cue-controlled reward choice: Effects of Pavlovian extinction on outcome-selective Pavlovian-instrumental transfer. *Journal of Experimental Psychology: Animal Learning and Cognition*, *44*, 280–292. <https://doi.org/10.1037/xan0000176>

Watson, P., Wiers, R. W., Hommel, B., & de Wit, S. (2014). Working for food you don't desire. Cues interfere with goal-directed food-seeking. *Appetite*, *79*, 139–148.

<https://doi.org/10.1016/j.appet.2014.04.005>

Watson, P., Wiers, R. W., Hommel, B., & de Wit, S. (2018). Motivational sensitivity of outcome-response priming: Experimental research and theoretical models. *Psychonomic Bulletin & Review*, 1–14. <https://doi.org/10.3758/s13423-018-1449-2>

Watson, P., Wiers, R. W., Hommel, B., Gerdes, V. E. A., & de Wit, D. (2017). Stimulus control over action for food in obese versus healthy-weight individuals. *Frontiers in Psychology*, *8*. <https://doi.org/10.3389/fpsyg.2017.00580>

Acknowledgements

We thank Diana Kviatkovskaja, James Nesaratnam, Adam Bennet, and Megan Barrett for their help with data collection.

Table 1*Experimental design*

Instrumental training	Pavlovian training	Instrumental booster	Outcome devaluation	Transfer test
R1 – O1, O3	S1 – O1	R1 – O1, O3	O1+, O3-	S1: R1 vs R2?
R2 – O2, O4	S2 – O2	R2 – O2, O4	O2+, O4-	S2: R1 vs R2?
	S3 – O3			S3: R1 vs R2?
	S4 – O4			S4: R1 vs R2?

Note: R1 and R2 denote instrumental responses (“E” and “I” key presses), O1 to O4 denote outcomes (crisps, popcorn, cashews and nachos points), and S1 to S4 denote Pavlovian stimuli (coloured squares). Plus (+) signs represent valued outcomes; minus (-) signs represent devalued outcomes.

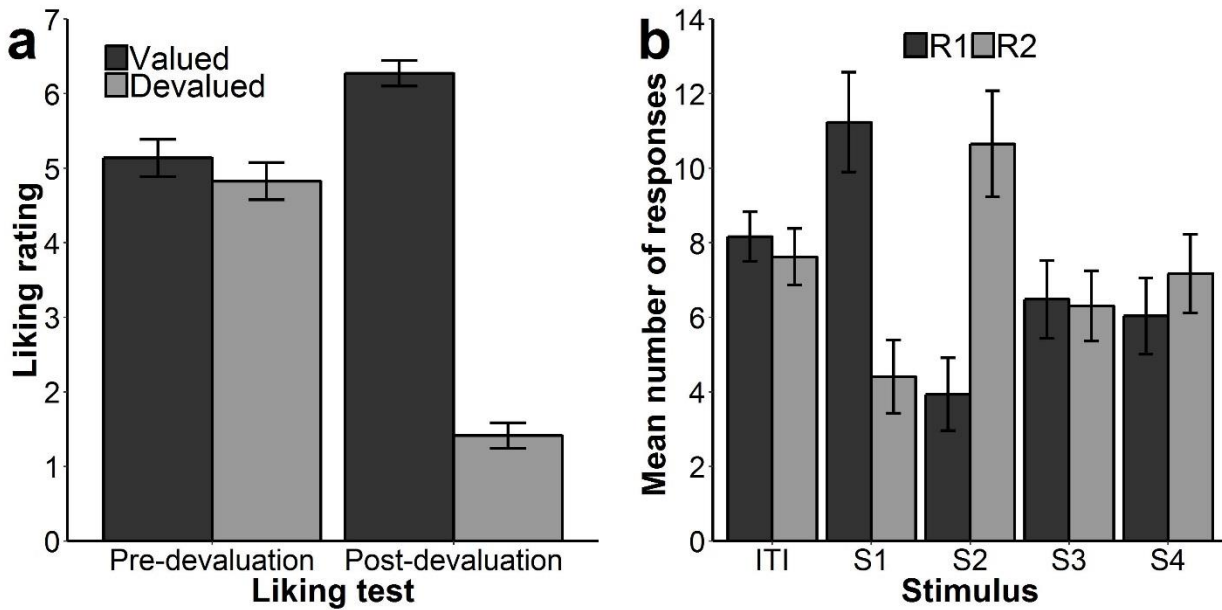


Figure 1. (a) Mean liking ratings for the valued (O1 and O2) and devalued (O3 and O4) outcomes at the start of the experiment (pre-devaluation) and immediately after the devaluation procedure (post-devaluation). Ratings of 1 and 7 represent wanting to eat the outcome “not at all” and “very much”, respectively. **(b)** Transfer test results. Instrumental responding was assessed in extinction in the presence of stimuli (S1-S4) that signalled each outcome, relative to inter-trial interval (ITI) periods. Error bars are difference-adjusted within-subject 95% confidence intervals (Baguley, 2012).