

## BRIEF COMMUNICATION OPEN



# Deliberately making mistakes: Behavioural consistency under win maximization and loss maximization conditions

Yajing Zhang<sup>1</sup>✉, Thi Kim Truc Huynh<sup>1</sup> and Benjamin James Dyson<sup>1,2</sup>✉

We argue that the feedback traditionally used to indicate negative outcomes causes future detrimental performance because of the default goal of *win maximization*. In gaming paradigms where participants intentionally performed as well (*win maximization*) and as poorly (*loss maximization*) as possible, we showed a double dissociation where actions following wins were more consistent during *win maximization*, but actions following losses were more consistent during *loss maximization*. This broader distinction between *goal-congruent* and *goal-incongruent* feedback suggests that individuals are able to flexibly redefine their definition of 'success', and provide a reconsideration of the way we think about 'losing'.

npj Science of Learning (2023)8:55; <https://doi.org/10.1038/s41539-023-00206-6>

How did you react to the error in our title “*Deliberately making mistakes*”? If you assumed that this was unintentional, then your reaction was probably an unfavourably negative one: we have been careless in our proofreading. However, in learning that the error was intentional, your reaction loses some of its negative connotations. This example clearly demonstrates that there is an interaction between the properties of a stimulus and the goal behind stimulus production<sup>1,2</sup>. Orthographically, the printed word “mistakes” remains a spelling error, independently of whether the spelling was unintentional or intentional. However, the negativity associated with this error can be reconciled with the knowledge of intentionally incorrect spelling. In this paper, we empirically investigate the interaction between the positive or negative nature of a stimulus, and, whether the production of that stimulus is incongruent or congruent with the goal set by the individual.

A dominant goal in most organismic behaviour is *win maximization*, typically defined as gaining as many/much as one can<sup>3–5</sup>. Both experimental and everyday tasks are similarly framed in terms of doing one's best: a memory test might require participants to try to remember in as much detail as possible<sup>6</sup>, a perceptual test might insist that performance is both as fast and as accurate as possible<sup>7</sup>, and, an individual visiting a casino hopes to leave with as much money as possible. Therefore, with behaviour oriented towards the goal of *win maximization*, it should not be surprising that when participants receive feedback that a memory detail is inaccurate, a perceptual error has been made, or money is lost, the delivery of this information can have negative consequences<sup>8–10</sup>. Furthermore, *Operant Conditioning*<sup>11,12</sup> dictates that there are very clear consequences for future actions on the basis of current outcomes: organisms will tend to repeat actions after a win (*win-stay*) but change actions after a loss (*lose-shift*). Although *lose-shift* behaviour is sub-optimal, it remains more predictable than *win-stay*<sup>11</sup> and also becomes more likely as the cognitive load of the task increases<sup>12</sup>. By these accounts, we would do well to avoid losses<sup>13</sup>.

However, a reimagining of the consequences of putatively negative feedback is possible by considering the role of goal-directed learning. When your goal is *win maximization*, on any trial that you are informed that you lost there is a discrepancy between your goal state and your current state (consistency theory; see

refs. <sup>14,15</sup>). This discrepancy we will refer to as *goal-incongruent* feedback. By this logic however, a second example of discrepancy presumably arises under the unique case when the goal of losing is violated by the experience of winning. To be specific, if *win maximization* is the goal, then each individual win represents *goal-congruent* feedback, whereas each individual loss represents *goal-incongruent* feedback. Conversely, if *loss maximization* is the goal, each individual loss represents *goal-congruent* feedback, whereas each individual win represents *goal-incongruent* feedback.

Therefore, much of the data accrued on the negative impact of losing may simply be due to the implicit (or explicit) goal of *win maximization*. If reactions to wins and losses are essentially arbitrary and individuals are able to flexibly redefine their definition of success, we should also see wins disrupting behaviour when the organism aims to lose. We carried out four experiments in which participants competed against computerized opponents under conditions of *win-maximization* and *loss-maximization*. In all experiments, different forms of computerized strategy were designed such that optimized participant behavior was either aligned or misaligned with the operant conditioning principles of *win-stay* and *lose-shift*. In Experiment 1, participants played *Matching Pennies* with the goal of *win-maximization*, where ‘wins’ represented *goal-congruent* and ‘losses’ represented *goal-incongruent* feedback. Our data showed that behavioral consistency was disrupted when goal and feedback were incongruent: specifically, performance was worse following losses relative to wins when the goal was *win maximization*. In Experiment 2, participants played the same game with the explicit goal of *loss-maximization*. Again, consistency was disrupted when the goal and feedback were incongruent but the behavioural observation was reversed: performance was worse following wins relative to losses when the goal was *loss maximization*. In Experiment 3, we confirmed the between-participant observations of Experiments 1 and 2 by running a within-participants design. In Experiment 4, we extended these findings to a different game (*Dice Dual*) involving 6 rather than 2 responses. Across Experiments 1–4, we also found no evidence that behaviour was significantly optimized when participant strategy was aligned with operant conditioning principles. Thus, our data demonstrate a double dissociation where behaviour following wins was more consistent during *win*

<sup>1</sup>University of Alberta, Edmonton, Canada. <sup>2</sup>Toronto Metropolitan University, Toronto, Canada. ✉email: [yajing10@ualberta.ca](mailto:yajing10@ualberta.ca); [bjdyson@ualberta.ca](mailto:bjdyson@ualberta.ca)

**Table 1.** Opponency structure and optimal strategies in Experiments 1–4 where [H] = Heads [T] = Tails [O] = Odd number [E] = Even number.

Opponent	Optimal Strategy		
		Win Maximization (Exp. 1, 3)	Loss Maximization (Exp. 2, 3)
<i>Unexploitable</i>			
[HHHHH] + [TTTTT] (random)	x 9	n/a	n/a
<i>Exploitable via repetition</i>			
[HHHHH] + [TTTTT] (random)	x 3	n/a	n/a
[HHHHHHHHHHH]	x 3	Win-Stay / Lose-Shift	Win-Shift / Lose-Stay
[TTTTTTTTTTT]	x 3	Win-Stay / Lose-Shift	Win-Shift / Lose-Stay
<i>Exploitable via alternation</i>			
[HHHHH] + [TTTTT] (random)	x 3	n/a	n/a
[HTHTHTHTHT]	x 3	Win-Shift / Lose-Stay	Win-Stay / Lose-Shift
[THTHTHTHTH]	x 3	Win-Shift / Lose-Stay	Win-Stay / Lose-Shift
		Win Maximization (Exp. 4)	Loss Maximization (Exp. 4)
<i>Unexploitable</i>			
[OOOOO] + [EEEE] (random)	x 9	n/a	n/a
<i>Exploitable via repetition</i>			
[OOOOO] + [EEEE] (random)	x 3	n/a	n/a
[OOOOOOOOO]	x 3	Win-Stay / Lose-Shift	Win-Shift / Lose-Stay
[EEEEEEEEEE]	x 3	Win-Stay / Lose-Shift	Win-Shift / Lose-Stay
<i>Exploitable via alternation</i>			
[OOOOO] + [EEEE] (random)	x 3	n/a	n/a
[OEOEOEOEO]	x 3	Win-Shift / Lose-Stay	Win-Stay / Lose-Shift
[EOEOEOEOE]	x 3	Win-Shift / Lose-Stay	Win-Stay / Lose-Shift

**Table 2.** Descriptive statistics of win rates and degrees of *win-stay* and *lose-shift* across Experiments 1–4 Standard error in parenthesis.

	Win Maximization			Loss Maximization		
	Win Rate	Win-stay	Lose-shift	Win Rate	Win-stay	Lose-shift
Experiment 1						
<i>Exploitable via repetition</i>	0.657 (.013)	0.800 (.018)	0.557(.020)		n/a	
<i>Exploitable via alternation</i>	0.601 (.015)	0.370 (.032)	0.381 (.016)			
Experiment 2						
<i>Exploitable via repetition</i>		n/a		0.369 (.018)	0.452 (.029)	0.228 (.032)
<i>Exploitable via alternation</i>				0.420 (.018)	0.569 (.026)	0.609 (.036)
Experiment 3						
<i>Exploitable via repetition</i>	0.655 (.013)	0.787 (.019)	0.557 (.015)	0.394 (.017)	0.535 (.023)	0.239 (.023)
<i>Exploitable via alternation</i>	0.603 (.014)	0.342 (.029)	0.431 (.017)	0.420 (.014)	0.615 (.020)	0.576 (.029)
Experiment 4						
<i>Exploitable via repetition</i>	0.607 (.014)	0.724 (.022)	0.524 (.019)	0.416 (.014)	0.521 (.021)	0.311 (.025)
<i>Exploitable via alternation</i>	0.583 (.013)	0.383 (.027)	0.427 (.019)	0.426 (.014)	0.288 (.018)	0.588 (.031)

*maximization*, but behaviour following losses was more consistent during *loss maximization*.

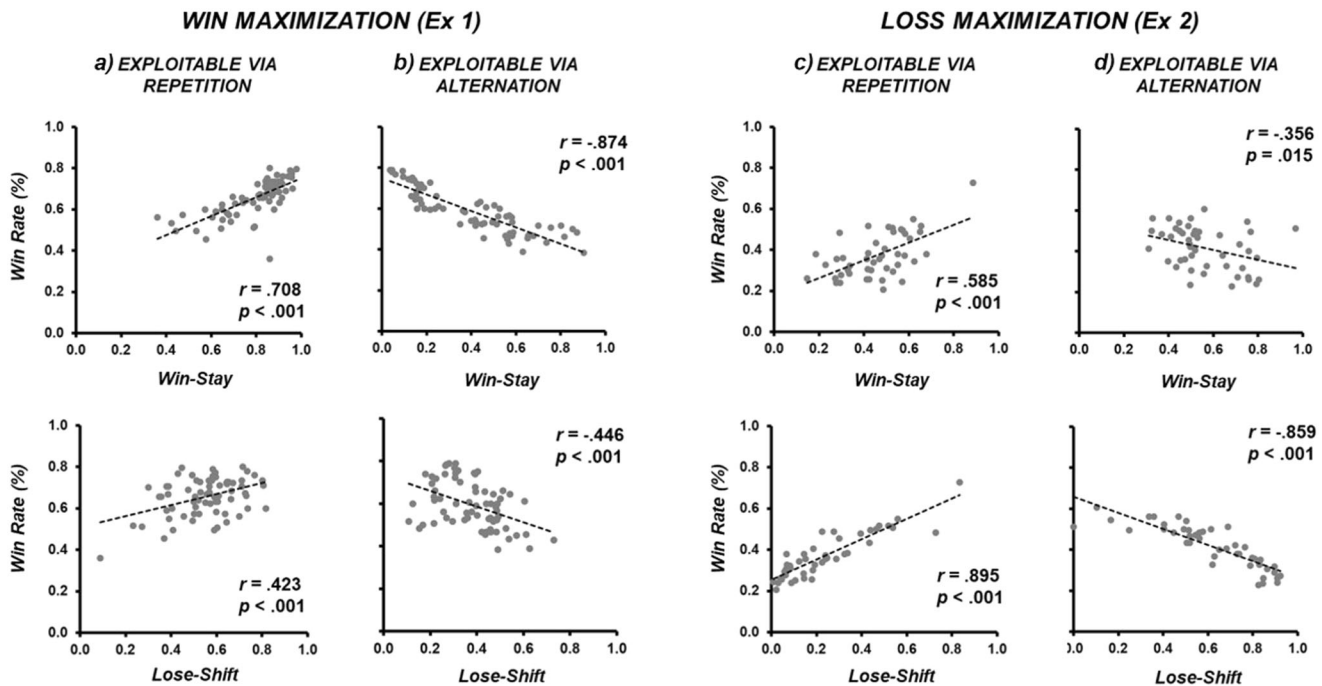
## RESULTS

### Measuring behavioural consistency

Across the experimental series, we test the idea that individuals are able to flexibly redefine their definition of success by measuring behavioural consistency when playing a game intentionally both as well as possible (i.e., *win maximization*) and as badly as possible (i.e., *loss maximization*). During *win maximization*, putatively positive feedback ('win') represents

*goal-congruency*. Conversely, negative feedback ('lose') represents *goal-congruency* during *loss maximization*. If receiving negative feedback is inherently detrimental, then performance should be universally less consistent following losses relative to wins. If, instead, behavioural disruption is determined by *goal incongruency*, performance should be less consistent following *loss* trials within the context of *win maximization* but also less consistent following *win* trials within the context of *loss maximization*.

Across Experiments 1–4, the strategy required for exploitation was determined by both the nature of opponency (*repetition*, *alternation*) and maximization goal (*win*, *loss*). As summarized in Table 1, in the case of *win maximization* for the *repetition*



**Fig. 1** Scatterplots of *win-stay* and *lose-shift* behaviour as a function of individual win rate, and the nature of goals and opponency. A double dissociation is observed between Experiments 1 and 2. **a, b** describe results in Experiment 1: the degree of behavioural consistency is significantly greater following a winning rather than losing outcome when the goal is *win maximization*; (**c, d**) describe results from Experiment 2: the degree of behavioural consistency is significantly greater following a losing rather than winning outcome when the goal is *loss maximization*. In both Experiments 1 and 2, this is irrespective of whether the action is consistent or inconsistent with the fundamental reinforcement learning principles of *win-stay* and *lose-shift*.

opponent, goal-consistent performance was guaranteed via the expression of traditional *win-stay* and *lose-shift* mechanisms: participants repeat their previous response following a win but change their previous response following a loss. However, *win maximization* for the *alternation* opponent could only be achieved by expressing the opposing rules: *win-shift* and *lose-stay*. These assignments were reversed when the goal was *loss maximization*: *win-shift* and *lose-stay* were required to lose against the *repetition* opponent, and, *win-stay* and *lose-shift* were required to reliably lose against the *alternation* opponent. As a result of this design, we were also able to assess the ‘default’ nature of actions traditionally associated with wins and losses as dictated by operant conditioning (*win-stay* / *lose-shift*; see<sup>16,17</sup>). If there is hardwired precedence for *win-stay* and *lose-shift*, then performance should be more consistent when strategy aligns with these principles (specifically, *win maximization* + *repetition* opponent, and, *loss maximization* + *alternation* opponent). However, if these outcome-action associations are arbitrary, consistency in expressing *win-shift* should be similar to *win-stay*, as should consistency in expressing *lose-stay* relative to *lose-shift*.

### Experiment 1 (win maximization)

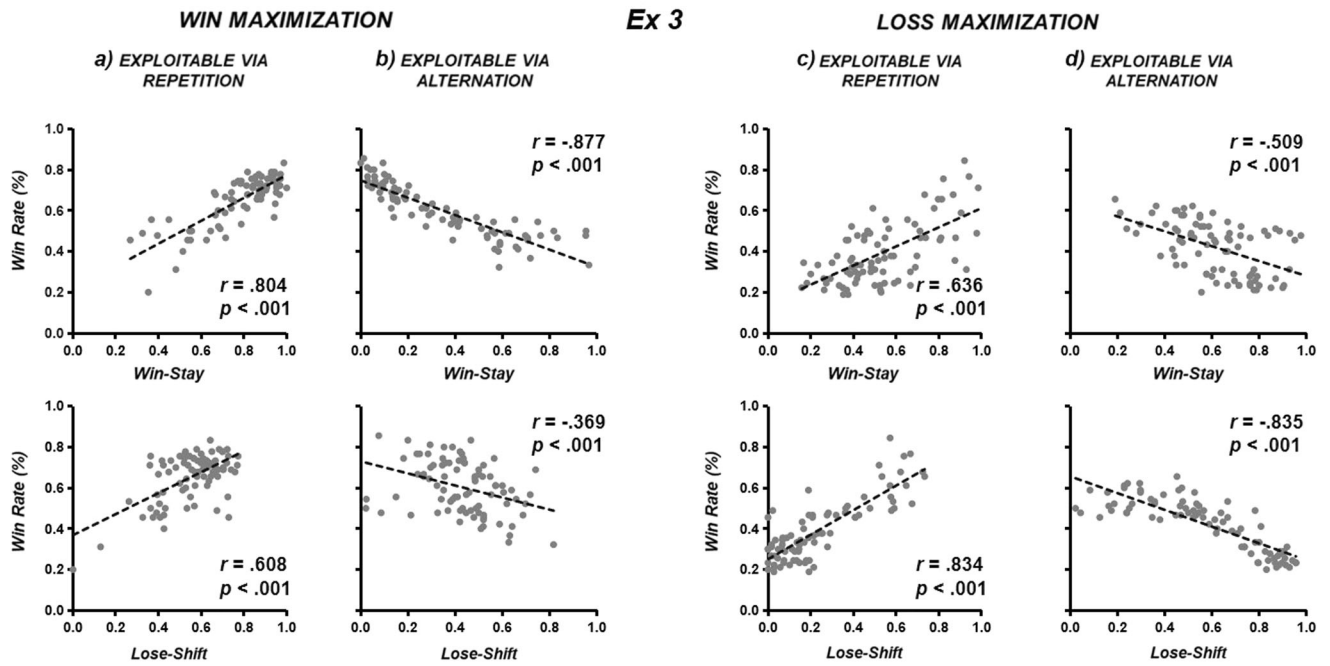
Win rates (proportion of a participant’s winning trials over all trials), the degree of *win-stay* (a proportion of *win-stay* trials over the total number of winning trials), and the degree of *lose-shift* (a proportion of *lose-shift* trials over the total number of losing trials) are calculated across Experiments 1–4 (see Table 2). During *win maximization* against an opponent *exploitable via repetition* (Experiment 1; Fig. 1a), both the degree of *win-stay* ( $r = 0.708$ ,  $p < 0.001$ ) and *lose-shift* ( $r = 0.423$ ,  $p < 0.001$ ) were positively correlated with win rate. By comparing the absolute values of the abovementioned correlation coefficients (i.e.,  $r = 0.708$  and  $r = 0.423$ ) using a z-test (two-tailed z test; see<sup>18,19</sup>), we found that stay actions following *wins* had a significantly stronger correlation

than shift actions following *losses* ( $z = 2.347$ ,  $p = 0.019$ ). Therefore, performance was more consistent following wins relative to losses in the context of *win maximization*.

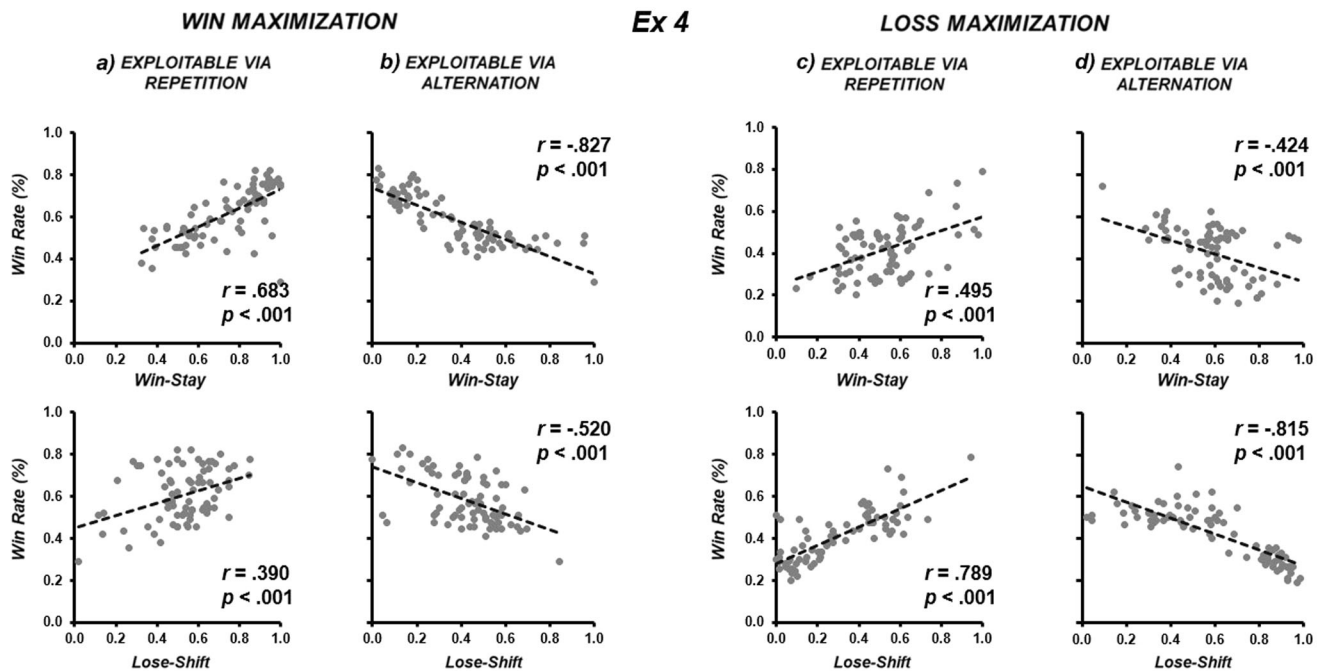
With the same goal of *win maximization* against an opponent *exploitable via alternation* (Fig. 1b), both the degree of *win-stay* ( $r = -0.874$ ,  $p < 0.001$ ) and *lose-shift* ( $r = -0.446$ ,  $p < 0.001$ ) were negatively correlated with win rate. This is consistent with the requirement to *win-shift* and *lose-stay* (see Method). Actions following *wins* once again had a significantly stronger correlation than actions following *losses* ( $z = 5.019$ ,  $p < 0.001$ ). Therefore, performance was more consistent following wins relative to losses in the context of *win maximization*. Thus, the data from Experiment 1 are clear that the requirement to deploy either *stay* or *shift* behaviour as a function of opponent was more consistent following *wins* relative to *losses*, when the goal of the task is *win maximization*.

### Experiment 2 (loss maximization)

During *loss maximization* against an opponent *exploitable via repetition* (Experiment 2; Fig. 1c), both *win-stay* ( $r = 0.585$ ,  $p < 0.001$ ) and *lose-shift* ( $r = 0.895$ ,  $p < 0.001$ ) behaviour were again positively correlated with win rate. In contrast to Experiment 1, actions following *losses* now had a significantly stronger correlation than actions following *wins* ( $z = 3.895$ ,  $p < 0.001$ ). Therefore, performance was more consistent following losses relative to wins in the context of *loss maximization*. Against an opponent *exploitable via alternation* (Fig. 1d), both the degree of *win-stay* ( $r = -0.356$ ,  $p = 0.015$ ) and *lose-shift* ( $r = -0.859$ ,  $p < 0.001$ ) were negatively correlated with win rate. This is consistent with the requirement to *win-shift* and *lose-stay* (see Method). Actions following *losses* had a significantly stronger correlation than actions following *wins* ( $z = 4.057$ ,  $p < 0.001$ ), showing that performance was more consistent following losses relative to wins in the context of *loss*



**Fig. 2** The double dissociation between Experiments 1 and 2 replicates in the within-participants design of Experiment 3. **a, b** Both *win-stay* and *win-shift* are more consistent in the context of *win maximization*, however, **(c, d)** both *lose-stay* and *lose-shift* are more consistent in the context of *loss maximization* (right four panels).



**Fig. 3** The double dissociation between Experiments 1 and 2 replicates in the within-participants design of Experiment 4. **a, b** Both *win-stay* and *win-shift* are more consistent in the context of *win maximization*, however, **(c, d)** both *lose-stay* and *lose-shift* are more consistent in the context of *loss maximization* (right four panels).

*maximization*. Collectively, Experiments 1 and 2 represent a double dissociation. When the goal of the task was *win maximization*, the deployment of either *stay* or *shift* behaviour was more consistent following wins relative to losses, but when the goal of the task was *loss maximization*, the deployment of either *stay* or *shift* behaviour was more consistent following losses relative to wins.

#### Experiments 3 and 4 (*win* and *loss maximization*)

This same double dissociation observed in the within-participants design of Experiment 3 (again using the 2-response game *Matching Pennies*; see Fig. 2) was similarly replicated in the within-participants design of Experiment 4 (using the 6-response game *Dice Dual*; see Fig. 3). That is, the behaviour was more consistent both following wins relative to losses in the

context of *win maximization* but also more consistent following losses relative to wins in the context of *loss maximization*.

In the context of *win maximization* against an opponent *exploitable via repetition*, the ability to enact *win-stay* behaviour was more consistent than the ability to enact *lose-shift* behaviour (Experiment 3:  $z = 2.744$ ,  $p = 0.006$ ; Fig. 2a; Experiment 4:  $z = 2.370$ ,  $p = 0.018$ ; Fig. 3a). In the context of *win maximization* against opponent *exploitable via alternation*, the negative correlation between win rate and *win-stay* behaviour was more consistent than *lose-shift* behaviour (Experiment 3:  $z = 6.047$ ,  $p < 0.001$ ; Fig. 2b; Experiment 4:  $z = 3.6676$ ,  $p < 0.001$ ; Fig. 3b). In the context of *loss maximization* against opponent *exploitable via repetition*, the ability to enact *lose-shift* behaviour was more consistent than the ability to enact *win-stay* behaviour (Experiment 3:  $z = 2.250$ ,  $p = 0.024$ ; Fig. 2c; Experiment 4:  $z = 3.1041$ ,  $p = 0.002$ ; Fig. 3c). In the context of *loss maximization* against opponent *exploitable via alternation*, the negative correlation between win rate and *lose-shift* was more consistent than *win-stay* behaviour (Experiment 3:  $z = 4.060$ ,  $p < 0.001$ ; Fig. 2d; Experiment 4:  $z = 3.9443$ ,  $p < 0.001$ ; Fig. 3d).

### Comparing win-stay / lose-shift against win-shift / lose-stay

Furthermore, we observed no evidence that *win-stay/lose-shift* were privileged forms of outcome-action association. In the context of *win maximization* in Experiment 1, participants were significantly less consistent in their ability to deploy *win-stay* behaviour (during *repetition*) than *win-shift* behaviour (during *alternation*) to intentionally increase win rate ( $r = 0.708$  vs [abs]  $r = 0.874$ ;  $z = -2.935$ ,  $p = 0.003$ ; two top left panels of Fig. 1). In the context of *loss maximization* (Experiment 2), participants were no less consistent in their ability to deploy *lose-stay* behaviour (during *repetition*) than *lose-shift* behaviour (during *alternation*) to intentionally decrease win rate ([abs]  $r = 0.895$  vs.  $r = 0.859$ ;  $z = 1.283$ ,  $p = 0.199$ ; two bottom right panels of Fig. 1). In Experiment 3, *win-stay* behaviour (during *repetition*) was numerically less consistent than *win-shift* behaviour (during *alternation*) during *win maximization* ( $r = 0.804$  vs [abs]  $r = 0.877$ ;  $z = -1.601$ ,  $p = 0.109$ ; two top left panels of Fig. 2). No difference was shown between *lose-shift* behaviour (during *alternation*) and *lose-stay* behaviour (during *repetition*) during *loss maximization* ([abs]  $r = 0.835$  vs.  $r = 0.834$ ;  $z = 1.086$ ,  $p = 0.278$ ; two bottom right panels of Fig. 2). In Experiment 4, *win-stay* behaviour (during *repetition*) was significantly less consistent than *win-shift* behaviour (during *alternation*) during *win maximization* ( $r = 0.683$  vs [abs]  $r = 0.827$ ;  $z = -1.9632$ ,  $p = 0.0496$ ; two top left panels of Fig. 3). No difference was shown between *lose-shift* behaviour (during *alternation*) and *lose-stay* behaviour (during *repetition*) during *loss maximization* ([abs]  $r = 0.815$  vs.  $r = 0.789$ ;  $z = 0.4173$ ,  $p = 0.677$ ; two bottom right panels of Fig. 3).

As an additional test of whether *win-stay / lose-shift* represented privileged forms of outcome-action associations, we also examined reaction times to see whether the speed to initiate these specific responses was faster than their reversed counterparts: *win-shift / lose-stay*. However, across Experiments 1–4, we found no significant interaction between outcome and action, such that *win-stay* and *lose-shift* timings were not significantly faster than *win-shift* and *lose-stay* timings (see Supplement Materials). Therefore, increased behavioural consistency following wins when the goal-state was winning (and following losses when the goal-state was losing) was independent of whether the individual is required to repeat (*stay*) or change (*shift*) their response to maintain maximization.

## DISCUSSION

There are two major implications of this work. First, our data show that people are flexible in defining ‘success’ specifically via the

malleable interpretation of putatively positive (*win*) and negative (*lose*) forms of feedback. We suggest that the traditional disruption generated by losses is the result of incongruity with the goal (implicit or explicit) of *win maximization*<sup>20,21</sup>. In support of this idea, we similarly showed disruption generated by wins due to an equivalent incongruity between feedback and goal in the context of *loss maximization*. This shows the top-down control we have over the interpretation of feedback, and the clear interaction between the properties of a stimulus and the goal behind stimulus production<sup>1,2</sup>. Our observations that behaviour is similarly disrupted by losing in the context of pursuing wins, and, winning in the context of pursuing losses are also consistent with the framework of goal-directed learning<sup>22,23</sup>. In the current studies, when the goal was to maximize wins (Experiments 1, 3 and 4), there was little conflict between what people were asked to do and the goal of *win maximization* implicit in most daily interactions. Much more unusual was the request to adjust one’s goal towards *loss maximization* (Experiments 2, 3 and 4), and our interest was in whether the same behavioural inconsistency would be demonstrated when putatively positive feedback (*win*) became the source of goal incongruity. Although our data suggest flexibility in goal-directed learning, one caveat is the degree to which the sensory properties of the visually-presented words ‘win’ and ‘lose’ constitute intrinsically positive and negative signals. One reason for the flexibility we observed may be because the use of points system represents abstracted rather than concrete forms of feedback. Therefore, it will be of interest to see whether behaviour consistency remains isomorphic between *win maximization* and *loss maximization* conditions when more tangible rewards and punishments are associated.

Second, our data show a remarkable ease with which we can switch out of putatively fundamental operant conditional outcome-action associations represented by *win-stay* and *lose-shift*<sup>24</sup>. Our fluid adaptation to counter reinforcement-learning strategies such as *win-shift* is consistent with data from nectarivorous birds and other organisms who also flexibly adapt to environments that have high depletion rates<sup>25,26</sup>. However, humans may be unique in our ability to simulate the sabotaging of our own performance in pursuit of the goal of *loss maximization* demonstrated here. For examples within popular culture, see the board game *Go For Broke*<sup>27</sup>, wherein players must *lose* \$1 million dollars, and, the UK entertainer Les Dawson who delighted audiences with deliberately terrible piano playing. Moreover, we have shown that both of these findings are observed in both binary (Experiments 1–3) and non-binary (Experiment 4) decision-making spaces, thereby lending credence to the generalizability of these findings.

In conclusion, the historical emphasis on *win-stay* and *lose-shift* mechanisms, and the presumed disruption of performance as a result of losses, are due to the often unspoken goal of *win maximization* within empirical sciences. We have shown that incongruent performance generated by the experience of losing is simply the result of a mismatch between an expected goal state and the current observed state. The same mismatch *and the same consequences of that mismatch* are produced when the intention is to lose but instead the organism wins. The flexibility with which individuals can define ‘success’ and specifically the interpretation of negative feedback is consistent with other observations such as deliberately seeking loss as in the principle of ‘fun failure’<sup>28,29</sup> or demonstrating cognitive proficiency via the act of intentional self-sabotage. This highlights the importance of subjective interpretation in decision-making, and may help to explain the individual differences associated with emotional reactions to feedback. Finally, we may wish to re-evaluate the extent to which our understanding of human and animal behaviour sciences has been limited by our implicit focus on *win maximization*.

## METHOD

Data from convenience samples of 71 (Mean = 19.77, SD = 3.12, 33 female), 46 (Mean = 19.2, SD = 1.67, 23 female), 84 (Mean = 19.47, SD = 3.25, 40 female), and 77 (Mean = 20.84, SD = 6.12, 38 female) participants were analyzed from the student population at the University of Alberta for Experiments 1, 2, 3, and 4, respectively. These sample sizes exceed the 30–40 participants previously analysed for zero-sum games in our lab (see refs. <sup>13,30</sup>), and, the minimum sample size of 29 participants for each experiment calculated using G\*Power<sup>31</sup> with large treatment effect (i.e., correlation  $\rho$  H1 = 0.5), alpha value at 0.05, and power level at 0.8. All participants were provided with written informed consent. They gave their informed consent for inclusion before they participated in the study. They completed the study for course credit and the on-line protocol was approved by the University of Alberta Research Ethics Committee (Pro00102699 and Pro00112365). Paradigms were controlled by Presentation 20.2 (build 07.25.18) and delivered remotely after participants downloaded Presentation Package Player. Two exclusion criteria were implemented: (1) procedural: where a participant stopped and re-started the paradigm thereby completing at least one condition more than once, and (2) behavioural: where a participant selected the same item 100% of the time throughout at least one condition. 8 (Experiment 1), 8 (Experiment 2), 22 (Experiment 3), and 6 (Experiment 4) participants were excluded from analyses according to the procedural criterion, and a further 4 (Experiment 1), 3 (Experiment 2), 8 (Experiment 3), and 2 (Experiment 4) participants were excluded from analyses according to the behavioural criterion. Data patterns across Experiments 1–4 remain consistent when participants excluded due to the behavioural criteria were added to the sample.

### Experiments 1–3

For Experiments 1–3, participants played 540 rounds of *Matching Pennies* consisting of 6 conditions each containing 90 trials. Each trial, both participant and computerized opponent selected either Heads or Tails. Participants won the trial if coin sides mismatched and lost the trial if coin sides matched. The 6 counterbalanced conditions consisted of the presence or absence of a cumulative score, crossed with three different kinds of opponency (*unexploitable*, *exploitable via repetition*, *exploitable via alternation*; see Table 1). Our factor pertaining to cumulative score yielded no notable effects, so we collapse across score manipulation.

The 90 trials per condition were subdivided into 9 groups of 10 trials each, with groups randomized within conditions. For the *unexploitable* opponent, each group consisted of 5 Heads and 5 Tails response which were further randomized within each group (e.g., THTHHHTHT). For the *exploitable via repetition* opponent, 3 blocks were identical to the *unexploitable* opponent, 3 blocks consisted of 10 presentations of Heads in a row (e.g., HHHHHHHHHH), and, 3 blocks consisted of 10 presentations of Tails in a row (e.g., TTTTTTTTTT). For the *exploitable via alternation* opponent, 3 blocks were once again identical to the *unexploitable* opponent, 3 blocks consisted of 10 coin alternations beginning with a Head (e.g., HTHHTHTHT), and, 3 blocks consisted of 10 coin alternations beginning with a Tail (e.g., THTHTHTHT). Consequently, all opponents played an equal percentage (50%) of Heads and Tails within each condition. The unexploitable conditions served as a filler task between two exploitable conditions. Because there were no optimal strategies to be employed in unexploitable condition to maximize either wins or losses, we did not include the results from the condition in our analyses and focus exclusively on exploitable opponency.

At each trial, participants would press one of the two buttons corresponding to Heads [K] or Tails [L] prompted by a fixation cross. Both participant and opponent selections were shown on the left and right side of screen, respectively, for 1000 ms.

Selections were removed during a 500 ms pause, followed by either “WIN (+1)” or “LOSE (-1)” in green and red font, respectively, for 1000 ms. Scores were updated and the fixation cross returned.

In Experiment 1, the goal was *win maximization*. Before *win maximization* conditions, participants were instructed to “Just try to do as well as you can!” In Experiment 2, the goal was *loss maximization*. Before *loss maximization* conditions, participants were instructed to “Just try to do as BADLY as you can! Try to get the most negative score.” In Experiment 3, both *win maximization* (cf., Experiment 1) and *loss maximization* (cf., Experiment 2) goals were completed in a within-participants design across 6 counter-balanced conditions, without the cumulative score manipulation.

### Experiment 4

To examine whether the data exhibited for the binary-response game *Matching Pennies* generalized to non-binary paradigms, participants played 540 rounds of the lab-designed game *Dice Dual* in Experiment 4. Here, participants and computerized opponents chose one number from the six sides of a die using 6 linearly organized keys. Participants won the trial if the sum of the two sides was even and lost the trial if the sum was odd. Thus, *Dice Dual* was structurally isomorphic to *Matching Pennies* (the actual odd or even die side was irrelevant) but had surface differences with respect to the number of responses available. For the *unexploitable* opponent, each group randomized 5 odd numbers and 5 even numbers (e.g., 1563224516). For the *exploitable via repetition* opponent, 3 blocks were identical to the *unexploitable* opponent, 3 blocks consisted of 10 presentations of odd numbers in a row (e.g., 5331311551), and, 3 blocks consisted of 10 presentations of even numbers in a row (e.g., 4622442466). For the *exploitable via alternation* opponent, 3 blocks were once again identical to the *unexploitable* opponent, 3 blocks consisted of 10 number alternations beginning with an odd number (e.g., 5416321256), and, 3 blocks consisted of 10 number alternations beginning with an even number (e.g., 2152361456). Consequently, all opponent types played an equal percentage (50%) of odd and even numbers within each condition. In Experiment 4, both *win maximization* and *loss maximization* goals were completed in a counterbalanced, within-participants design (as per Experiment 3).

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

### DATA AVAILABILITY

All data is available in the [Supplement Materials](#).

Received: 7 September 2022; Accepted: 17 November 2023;

Published online: 06 December 2023

### REFERENCES

- Gallant, S. N., Dyson, B. J. & Yang, L. Local context effects during emotional item directed forgetting in younger and older adults. *Memory* **25**, 1129–1138 (2017).
- Awh, E., Belopolsky, A. V. & Theeuwes, J. Top-down versus bottom-up attentional control: a failed theoretical dichotomy. *Trends Cogn. Sci.* **16**, 437–443 (2012).
- Mookherjee, D. & Sopher, B. Learning behavior in an experimental matching pennies game. *Games Econ. Behav.* **7**, 62–91 (1994).
- Dechenaux, E., Kovenock, D. & Sheremeta, R. M. A survey of experimental research on contests, all-pay auctions, and tournaments. *Exp. Econ.* **18**, 609–669 (2015).
- Van Dijk, E. & De Dreu, C. K. W. Experimental games and social decision-making. *Annu. Rev. Psychol.* **72**, 415–438 (2021).
- Bousfield, W. A. & Rosner, S. R. Free vs uninhibited recall. *Psychon. Sci.* **20**, 75–76 (1970).

7. Mulder, M. J. et al. The speed and accuracy of perceptual decisions in a random-tone pitch task. *Atten. Percept. Psychophys.* **75**, 1048–1058 (2013).
8. Eben, C., Chen, Z., Vermeylen, L., Billieux, J. & Verbruggen, F. A direct and conceptual replication of post-loss speeding when gambling. *R. Soc. Open Sci.* **7**, 200090 (2020).
9. Eskreis-Winkler, L. & Fishbach, A. Not learning from failure- the greatest failure of all. *Psychol. Sci.* **30**, 1733–1744 (2019).
10. Torrance, J. et al. Conceptualising emotional and cognitive dysregulation amongst sports bettors; an exploratory study of ‘tilting’ in a new context. *PLoS One* **17**, e0264000 (2022).
11. Dyson, B. et al. Negative outcomes evoke cyclic irrational decisions in rock, paper, scissors. *Sci. Rep.* **6**, 20479 (2016).
12. Ivan, V. E., Banks, P. J., Goodfellow, K. & Gruber, A. J. Lose-shift responding in humans is promoted by increased cognitive load. *Front. Integr. Neurosci.* **12**, 9 (2018).
13. Metcalfe, J. Learning from errors. *Annu. Rev. Psychol.* **68**, 6.1–6.25 (2016).
14. Aronson, E. A theory of cognitive dissonance: A current perspective. In *Advances in Experimental Social Psychology* 2nd edn, Vol. 4 (eds L. Berkowitz) Ch. 1–34. (Academic Press, 1968).
15. Festinger, L. *A Theory of Cognitive Dissonance*. (Stanford University Press, 1957).
16. Thorndike, E. L. *Animal Intelligence*. (Macmillan Company, 1911).
17. Skinner, B. F. *Contingencies of Reinforcement*. (Appleton-Century-Crofts, 1969).
18. Diedenhofen, B. & Musch, J. Cocor: A comprehensive solution for the statistical comparison of correlations. *PLoS One* **10**, e0121945 (2015).
19. Dunn, O. J. & Clarke, V. Correlation coefficients measured on the same individuals. *J. Am. Stat. Assoc.* **64**, 366–377 (1969).
20. De Loof, E. et al. Signed reward prediction errors drive declarative learning. *PLoS One* **13**, e0189212 (2018).
21. Niv, Y. & Schoenbaum, G. Dialogues on prediction errors. *Trends Cogn. Sci.* **12**, 265–272 (2008).
22. Valentin, V. V., Dickinson, A. & O’Doherty, J. P. Determining the neural substrates of goal-directed learning in the human brain. *J. Neurosci.* **27**, 4019–4026 (2007).
23. Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P. & Dolan, R. J. Model-based influences on humans’ choices and striatal prediction errors. *Neuron* **69**, 1204–1215 (2011).
24. Sundvall, J. & Dyson, B. J. Breaking the bonds of reinforcement: effects of trial outcome, rule consistency and rule complexity against exploitable and unexploitable opponents. *PLoS One* **17**, e0262249 (2022).
25. Stagner, J. P., Michler, D. M., Rayburn-Reeves, R. M., Laude, J. R. & Zentall, T. R. Mid-session reversal learning: Why do pigeons anticipate and persevere? *Learn. Behav.* **41**, 54–60 (2013).
26. Sulikowski, D. & Burke, D. Shifting in nectarivorous birds: Selective inhibition of the learned win-stay responses. *Anim. Behav.* **83**, 519–524 (2012).
27. Bradley, M. *Go For Broke* [Board game] (Selchow & Righter Company, 1965).
28. McGonigal, J. *Reality is Broken: Why Games Make us Better and How They can Change the World* (Penguin Press New York, 2011).
29. Nguyen, C. T. *Games: Agency as Art* (Oxford University Press, 2020).
30. Dyson, B. J., Musgrave, C., Rowe, C. & Sandhur, R. Behavioural and neural interactions between objective and subjective performance in a Matching Pennies game. *Int. J. Psychophysiol.* **147**, 128–136 (2020).
31. Faul, F., Erdfelder, E., Buchner, A. & Lang, A.-G. Statistical power analyses using G\*Power 3.1: tests for correlation and regression analyses. *Behav. Res. Methods* **41**, 1149–1160 (2009).

## ACKNOWLEDGEMENTS

The studies were funded by an NSERC Discovery Grant (RGPIN-2019-04954), Alberta Gambling Research Grants, and start-up monies provided by the University of Alberta (RES0042096).

## AUTHOR CONTRIBUTIONS

B.J.D. designed the studies. Y.Z. and T.H.T.H. collected the data and initially analysed the data. Y.Z., T.H.T.H. and B.J.D. wrote the manuscript.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41539-023-00206-6>.

**Correspondence** and requests for materials should be addressed to Yajing Zhang or Benjamin James Dyson.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023