# scientific reports



# **OPEN** Modifying Pavlovian-to-instrumental transfer by approach avoidance training in healthy subjects: a proof of concept study

Annika Rosenthal<sup>1</sup>, Ke Chen<sup>1</sup>, Anne Beck<sup>2,4</sup> & Nina Romanczuk-Seiferth<sup>1,3,4</sup>

The modulation of instrumental action by conditioned Pavlovian cues is hypothesized to play a role in the emergence and maintenance of maladaptive behavior. The Pavlovian to Instrumental transfer task (PIT) is designed to examine the magnitude of the influence of cues on behavior and we aim to manipulate the motivational value of Pavlovian cues to reduce their effect on instrumental responding. To this end, we utilized a joystick-based modification of approach and avoidance propensities that has shown success in clinical populations. To examine changes in PIT, we subjected 35 healthy participants to a series of experimental procedures: (1) Instrumental training was followed by (2) Pavlovian conditioning of neutral stimuli that were associated with monetary reward or loss. (3) In a subsequent joystick task, approach and avoidance tendencies toward conditioned cues were assessed. (4) In a transfer test, the PIT effect as the impact of conditioned cues on instrumental behavior was measured. (5) The explicit knowledge of cue-reward contingencies was assessed in a forced-choice phase. (6, 7) systematic joystick training was followed by a posttest (8) the transfer task and forced-choice test were repeated. We found no effect of training on approach-avoidance propensities in the context of this proof of concept study. A higher response rate towards negative stimuli during PIT after systematic training compared to sham training was seen. On the other hand, we saw an increased PIT effect after sham training. These results contribute to the understanding of the strength of the influence of cues on instrumental behavior. Our findings further stress the importance of context, instructions and operationalization of instrumental behavior in the framework of transfer effects.

Stimuli that are associated with rewards have been demonstrated to encourage behaviors attributed to past rewarding experiences<sup>1,2</sup>. This so-called concept of cue-reactivity is central to human and animal adaptive behavior such as food seeking or reproduction<sup>2</sup>. Cue-reactivity entails that our actions are continuously influenced and guided by predictive cues, which can prompt or deter us from engaging in certain behaviors. As stated above, these cues can be either adaptive or maladaptive, leading to suboptimal choices. However, cue-reactivity has been mostly researched in the background of disparaging behavior such as substance abuse, binge eating, or other behavioral patterns marked by conflicts of behavioral goals and values assigned to stimuli<sup>3-5</sup>. For instance, after initially producing rewarding effects, prolonged drug abuse could alter motivational drive and sensitize to drug-related conditioned responding and craving<sup>67</sup>. These factors play a fundamental role in the maintenance of substance use disorders (SUDs). The interplay of instrumental behavior and reinforcing properties of stimuli has been conceptualized in various ways. One way cue-motivated behavior has been modeled, is through Pavlovianto-instrumental transfer (PIT) tasks<sup>8,9</sup>. The earliest studies on PIT can be traced back to the 1940s. These studies revealed that stimuli that were associated with food, could increase the likelihood of instrumental actions

<sup>1</sup>Department of Psychiatry and Neurosciences|CCM, Charité–Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany. <sup>2</sup>Faculty of Health, Health and Medical University, Campus Potsdam, Potsdam, Germany. <sup>3</sup>Department of Psychology, Clinical Psychology and Psychotherapy, MSB Medical School Berlin, Berlin, Germany. <sup>4</sup>These authors jointly supervised this work: Anne Beck and Nina Romanczuk-Seiferth. <sup>™</sup>email: annika.rosenthal@charite.de

directed towards food<sup>9</sup>. Transfer effects can have both facilitating and inhibiting effects on actions, as cues can either increase or decrease the frequency of certain actions or influence preference towards specific actions. The direction of transfer effects depends on several factors, such as the valence of the Pavlovian cue, which refers to whether it is appetitive or aversive. For instance, a Pavlovian cue linked to an aversive shock may encourage actions aimed at avoiding shock but discourage actions aimed at obtaining food (Rescorla and Solomon, 1967). This action-specific control of behavior has been shown and researched in humans as well<sup>10-12</sup>. Essentially, PIT tasks constitute instrumental training to establish response-outcome associations by linking responses to reward delivery. In a Pavlovian conditioning phase, previously neutral stimuli (CS) predict rewarding outcomes (US), to establish stimulus-outcome CS-US) associations. In the transfer test, instrumental behavior is assessed in extinction and in presence of the outcome-associated stimuli (CS-US)<sup>13</sup>. An increase of instrumental behavior due to high motivational salience (i.e., preferential attentional processing) of the Pavlovian CS has been found in various clinical populations as well as at-risk groups, i.e. alcohol use disorder<sup>14</sup>, social drinking<sup>15</sup>, obesity<sup>16</sup> as well as aversive PIT was exaggerated in patients with depression<sup>17</sup>, while some also found transfer effects to be reduced in patients with depression<sup>18</sup>. In addition, increased loss aversion PIT was found in subjects with gambling disorder<sup>19</sup>. Despite its' importance in psychopathology, the influence of environmental cues on behavior is pertinent to decision-making and instrumental choice behavior in general. Accordingly, PIT effects are a phenomenon that has been researched in healthy populations<sup>20</sup>. Nevertheless, the underlying mechanisms of the nature of transfer effects are debated<sup>21</sup>. Associative accounts postulate the dissociative engagement of motivational and cognitive control and within this context, biased action selection towards reward cues has been explained<sup>22</sup>. Similarly originating within a dual-process framework, that divides cognitive function into implicit and automatic or explicit and controlled processes, cognitive bias modification (CBM) was developed with the idea to evaluate these biases in the context of maladaptive behavior<sup>23</sup>. One form of CBM focuses on bias in the automatically activated action tendency to approach or avoid certain stimuli. This approach avoidance task (AAT) operationalizes push (avoid) and pull (approach) behavior with a joystick experiment in which subjects are instructed to react to content or content-unrelated features of a stimulus. Studies have shown that an increased approach bias towards drug-related stimuli was related to consumption or addiction severity<sup>24,25</sup> as well as food associated stimuli increased approach bias in food craving<sup>26</sup> and was correlated to uncontrolled eating<sup>27</sup>. In this context, modified versions of the AAT have been used to successfully retrain these altered approach/avoidance tendencies in clinical populations<sup>28-30</sup>. An indication of efficacy of AAT training is uncertain, however, due to mixed results for reviews, see<sup>31,32</sup>. In extension of this, Pavlovian conditioning of previously neutral cues has been shown to elicit significant approach tendencies towards these cues as well. For example, approach bias was enhanced when participants were faced with abstract stimuli that were previously paired with chocolate<sup>33</sup> or tobacco<sup>34</sup>.

Regarding the role of implicit motivational processes that drive behavior, PIT effects and approach bias have been theorized to influence each other but the exact nature of this has not been disentangled yet. Both concepts have been shown to overlap in salience attribution of external cues, ultimately driving behavior. This is known as e.g., incentive salience in the context of clinical disorders<sup>35</sup>. While it has been proposed that approach bias plays a role in PIT<sup>36</sup>, it has also been stipulated that transfer effects drive approach bias<sup>37</sup>. In addition, research suggests that the PIT and AAT are associated on a neuronal and behavioral level<sup>38</sup>. Considering clinical relevance, increased PIT effects in psychiatric populations, especially in the addiction domain, have not been subjected to systematic modification yet. It has been proposed, however, that clinical populations with strong PIT effects could profit from approach modification training to reduce PIT effects and increase behavioral control<sup>14</sup>. From a therapeutic perspective, it makes sense to systematically reduce the effect of Pavlovian cues that trigger maladaptive behavior: on the one hand, to prevent possible situational habit formation in at-risk populations and on the other hand, to disrupt the effect of environmental cues and ensure abstinence from already established dysfunctional behavior.

While the effect of alcohol-related AAT training on PIT effects has been investigated<sup>39</sup>, the direct manipulation of experimentally Pavlovian conditioned stimuli has not been attempted thus far. In light of this, we want to investigate whether PIT effects can be translated into approach and avoidance biases towards previously conditioned cues in the context of an AAT and subsequent retraining of these biases. To examine this in a healthy population, we will employ a previously established task that uses monetary cues<sup>40</sup>. We first want to evaluate whether Pavlovian conditioning of neutral cues in the context of a PIT task elicits approach and avoidance biases toward them. In a next step, we want to find out whether these action tendencies can be manipulated via a modified training version of the AAT. Lastly, we want to assess the effects of this systematic manipulation of the approach avoidance propensities of Pavlovian cues on Pavlovian-to-Instrumental transfer (PIT) processes in healthy adults.

# Methods

All procedures complied with the Declaration of Helsinki and were approved by the ethical committee of the Charité–Universitätsmedizin Berlin. All participants gave full written informed consent.

**Participants.** The study was conducted in Berlin, Germany and all participants were recruited through internet advertisement. Participants were included if between 18 and 65 years of age and they were neither pregnant nor breastfeeding. To exclude subjects with pathological manifestations of traits that could confound our results, exclusion criteria were personality disorders, lifetime bipolar disorder, acute depressive episode, and SUD (except tobacco use disorder (TUD) and mild (up to 6 criteria) cannabis use disorder (CUD)) according to the Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition) (DSM-5). The sample consisted of 35 participants (22 female) and age ranged from 19 to 60 (mean = 35.9 SD = 12.01) (*see* Table 1.).

|        | Sham training | Intervention training |         |
|--------|---------------|-----------------------|---------|
|        | Mean (SD)     | Mean (SD)             | $p/x^2$ |
| Age    | 37 (11.86)    | 35.15 (12.45)         | 0.66    |
|        | Male (female) | Male (female)         |         |
| Gender | 6 (9)         | 7 (13)                | 0.76    |

 Table 1. Descriptives of sham- and intervention training group. X<sup>2</sup> chi-squared, SD standard deviation.

.....

**Procedure.** The PIT paradigm was administered in six parts, which consisted of (1) instrumental training, (2) Pavlovian training, (3, 5) PIT before and after AA training, and (4, 6) a forced choice task before and after AA training (see Fig. 1A–F). The task was programmed with Matlab 2019b (MATLAB version 9.7.0, 2019; MathWorks, Natick, MA, USA) using the Psychophysics Toolbox Version 3 (PTB-3.0.15) extension<sup>41-43</sup>. For an extensive description of the PIT paradigm, please see Garbusow *et al.*<sup>40</sup>. After the Pavlovian conditioning phase, we administered the AAT to assess the participants' approach/avoidance bias. Following PIT and the forced choice task, the AAT training and post-test were administered (please see Fig. 1 for details of the study design).

**Instrumental training.** Instrumental stimuli consisted of six shells in various colors and shapes that were presented on a computer screen. The instructions to the participants were to collect "good" and leave "bad" shells while receiving probabilistic feedback. In order to collect a good shell, the subjects had to repeatedly press the left mouse button (at least 5 times), while they had to omit a reaction when a bad shell was presented (0–4 button presses were counted as omission in order to measure vigor). In a random fashion, correct responses were rewarded with 20 Cents in 80% of the trials and punished with a loss of 20 Cents in 20% of trials, and vice versa for incorrect responses. Dependent on performance, a learning criterion determined the task length to be between 60 and 120 trials (80% correct trials over 16 consecutive trials).

**Pavlovian conditioning.** Subjects were presented with 48 trials of abstract image sound combinations (compound CS) that were paired with monetary reward or punishment (US) in a deterministic fashion. The compound CS was presented for 3 s on the left or right side of the screen, followed by a delay of 3 s with two fixation crosses at the two potential CS locations, then a US (monetary reward 10€, punishment -10€ or no stimulus) was presented for another 3 s. The comparatively large amount of money is intended to increase the influence of Pavlovian conditioning, while the small amount is intended to elicit variable accuracy in the context of instrumental training. Subjects had to passively watch and memorize the CS and US pairings. Abstract pictures and associated US were randomly paired.



**Figure 1.** Schematic representation of the study design. Instrumental training (1) and Pavlovian conditioning (2) is followed by an AAT pretest (3). After PIT and Query trials (4,5), AAT training (6) and posttest (7) are followed by a second PIT and Query trials (8,9). AAT approach avoidance training, PIT Pavlovian-to-Instrumental transfer.

**Pavlovian-to-instrumental transfer.** The instrumental task was performed in formal extinction, which means that the monetary outcome was not shown. During the trials, the background was alternately tiled with one of the CS. The task had a duration of 108 trials with each lasting 3 s.

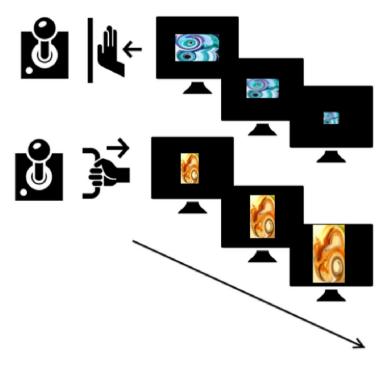
**Ouery trials.** To avoid mixing Pavlovian and instrumental conditioning, a forced choice task was administered after the transfer stage to confirm the success of the Pavlovian conditioning. Based on their subjective preference, subjects had to choose between two Pavlovian CSs (9 trials). All pairings were presented in an interleaved, randomized order.

**AAT.** This version of the AAT was programmed so that participants had to respond according to the orientation of the pictures (*see* Fig. 2). With a joystick, all horizontal pictures were to be pushed away (avoidance), and all vertical pictures were to be pulled closer (approach). To mimic realistic approach/avoidance behavior, a zoom feature was used. In the approach (pull) movement, pictures grew larger; while pushing resulted in shrinking pictures. The task structure was adapted from Wiers et al.<sup>44</sup>. As stimuli, we used the same abstract pictures (CS) as in the PIT task.

The AAT consisted of three different phases: a pre-test AAT, a training AAT, and a post-test AAT. A practice phase of 40 trials, in which participants learned to move the joystick according to orientation, was followed by 120 trials in which all three abstract pictures had to be pushed/pulled with equal frequency. During the training phase which consisted of 300 trials, participants moved the joystick according to the assigned condition. In the sham condition, positive and negative stimuli were shown equally often in push and pull trials. In the intervention condition, participants had to push positively conditioned pictures  $(10\varepsilon)$  and pull negatively conditioned pictures  $(-10\varepsilon)$  while neutral pictures were both pushed and pulled with equal frequency. In the post-test phase, 120 trials with equal frequency of all possible trial types were administered.

**Analysis.** *Statistical analysis plan.* Data were analyzed in Matlab 2011a and Jamovi (The Jamovi Project, 2022). Generalized linear modeling (GLM), as well as generalized linear mixed modeling (GLMM), was used since the RT data from the AAT were non-normally distributed (Kolmogorov–Smirnov: p < 0.001) and the PIT response showed a zero inflation as no response was required on some of the trials. All models used log link functions. The full fixed effects structure of the model was supplemented by the step-wise inclusion of random effects while testing their respective contribution to the model fit. Fixed effects and their interactions were tested using omnibus (Wald) Chi-Squared tests (results are reported in the supplementary Tables 1–4).

**AAT.** To exploratorily assess the individual contribution of Pavlovian CS on approach avoidance behavior at pre and post training, initial gamma distributed models were applied to test for the effect of background CS  $(-10\epsilon/neutral/+10\epsilon; dummy coded with - 10\epsilon as a reference)$ , direction (push/pull; coded as -0.5/+0.5) as well as their interaction on RT in each trial. Subsequently a GLMM, a gamma distributed model was built in which the RT in each trial was predicted by the value of the background CS  $(-10\epsilon/neutral/+10\epsilon; dummy coded with - 10\epsilon as a reference)$ .



**Figure 2.** In detail representation of the AAT training. Positively conditioned fractals are pushed away (top) and negatively conditioned fractals are pulled towards the subject (bottom). *AAT* approach avoidance training.

-10€ as a reference) the direction (push/pull; coded as -0.5/+0.5), the training condition (Sham/Avoidance; coded as -0.5/+0.5) and time (pre Training/post Training; coded as -0.5/+0.5) as well as their interactions. Initial model comparison of different random-effects structures indicated the best model fit for taking intercept, main effects of CS value, direction, and time as random effects across subjects. Please see supplementary Table 5 for model comparison based on the Akaike Information Criterion (AIC)<sup>45</sup> of different random-effects structures.

**PIT.** Initial model fitting indicated a Poisson distributed model to be overdispersed  $(Chi^2/DF=3.77)^{46}$ . To account for this, a model with negative binomial distribution was built to assess the individual contribution of Pavlovian values on behavior. Here, the number of button presses in each trial was predicted by the value of the background CS ( $-10\epsilon/neutral/+10\epsilon$ ; dummy coded with  $-10\epsilon$  as a reference) the instrumental condition (not collect/collect; coded as -0.5/+0.5), the training condition (Sham/Avoidance; coded as -0.5/+0.5) and time (pre Training/post Training; coded as -0.5/+0.5) as well as their interactions. Model comparisons indicated the within-subject factors intercept, main effect of CS value, and time to be the best fitting random factor structure. Please see supplementary Table 6 for model comparison based on AIC of different random-effects structures.

#### Results

Of the experimental trials, 6.92% were excluded (implausible RTs based on 1st and last percentile as well as all participants with commission errors > 21% based on<sup>47,48</sup>). The n = 1 excluded participant had error rates of 39.2% and 40% in the AAT pretest and posttest respectively.

**AAT.** In the pre training AAT, we did not find a significant effect of the background CS, direction, or their interaction (supplementary Table 7/Fig. 1).

For summary statistics of AAT RT pre and post training conditions, please see Table 2.

In our GLMM, the main effect of interest was the interaction between training by time by CS by direction on RT. This however, was not significant (*estimate* = 0.061; SE = 0.041; p = 0.140), indicating there was no effect of training on the change of approach or avoidance bias towards the CS (*please see* Table 3 and Fig. 3). However, there was a main effect of time (*estimate* = -0.075; SE = 0.025; p = 0.003), indicating a general decrease of RT from pre AAT to post AAT. Moreover, the direction by CS interaction was significant (*estimate* = 0.025; SE = 0.010; p = 0.016). Additionally, the CS by direction by time interaction was significant (*estimate* = 0.051; SE = 0.021; p = 0.015), indicating that irrespective of training, the RT of approach of positive CS was significantly higher compared to approach RT towards negative CS. Please see Fig. 3 for a graphic representation of the results of this interaction.

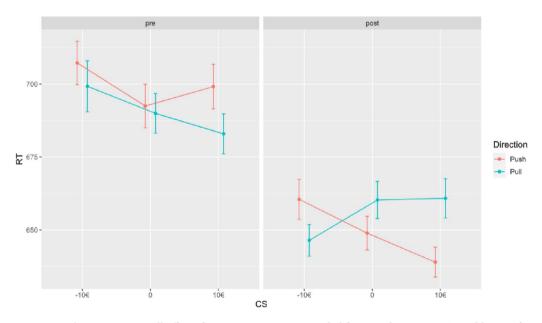
|           |      |                       |       | RT   |       |
|-----------|------|-----------------------|-------|--|-------|
| Direction | Time | Training condition    | CS    | Mean   | SE    |
|           |      |                       | - 10€ | 718.73   | 12.36 |
|           |      | Sham training         | 0     | 692.55   | 10.70 |
|           | Pre  |                       | 10€   | 708.87   | 11.89 |
|           | rie  |                       | - 10€ | 699.69   | 9.27  |
|           |      | Intervention training | 0     | 692.43   | 10.22 |
| Duch      |      |                       | 10€   | Mean         SE           10€         718.73         12.           692.55         10.           692.55         10.           0€         708.87         11.           10€         699.69         9.           692.43         10.           0€         692.75         9.           10€         692.75         9.           10€         692.75         9.           10€         659.85         8.           0€         654.71         7.           10€         659.62         9.           641.49         7.         7.           0€         628.06         6.           10€         688.45         12.           0€         682.31         9.           10€         689.62         9.           0€         683.35         9.           10€         648.66         8.           0€         665.07         10.           0€         659.27         9.           10€         644.90         6.           0€         659.97         7. | 9.95  |
| r usii    |      |                       | - 10€ | 661.74   | 9.08  |
| Push      |      | Sham training         | 0     | 659.85   | 8.78  |
|           | Post |                       | 10€   | 654.71   | 7.59  |
|           | 1031 |                       | - 10€ | 659.62   | 9.80  |
|           |      | Intervention training | 0     | 641.49   | 7.59  |
|           |      |                       | 10€   | 628.06   | 6.83  |
|           |      |                       | - 10€ | 688.45   | 12.21 |
|           |      | Sham training         | 0     | 690.46   | 10.30 |
|           | Pre  |                       | 10€   | 682.31   | 9.50  |
|           | 1 IC |                       | - 10€ | 706.13   | 12.01 |
|           |      | Intervention training | 0     | 689.62   | 9.03  |
| Dull      |      |                       | 10€   | 683.35   | 9.43  |
| ruii      |      |                       | - 10€ | 648.66   | 8.61  |
|           | Post | Sham training         | 0     | 666.07   | 10.76 |
|           |      |                       | 10€   | 659.27   | 9.87  |
|           |      |                       | - 10€ | 644.90   | 6.97  |
|           |      | Intervention training | 0     | 655.97   | 7.70  |
|           |      |                       | 10€   | 661.89   | 9.04  |

**Table 2.** Descriptives of reaction time in the approach avoidance task (AAT), stratified by direction, time, training condition, and CS. *RT* reaction time, *CS* conditioned stimulus, *SE* standard error.

Scientific Reports | (2023) 13:10074 |

| Fixed effects parameter estimates   |          |       |                            |         |         |         |         |  |
|---|----------|-------|----------------------------|---------|---------|---------|---------|--|
|   |          |       | 95% Confidence<br>Interval |         |         |         |         |  |
| Names   | Estimate | SE    | Lower                      | Upper   | exp(B)  | z       | р       |  |
| (Intercept)   | 6.502    | 0.031 | 6.441                      | 6.562   | 666.319 | 209.963 | < 0.001 |  |
| CS (10€–10€)  | - 0.007  | 0.011 | - 0.028                    | 0.013   | 0.993   | - 0.700 | 0.484   |  |
| Direction (pull)  | - 0.022  | 0.015 | - 0.051                    | 0.008   | 0.979   | - 1.449 | 0.147   |  |
| Time (post training)  | - 0.075  | 0.025 | - 0.124                    | - 0.026 | 0.928   | - 2.979 | 0.003   |  |
| Training (intervention training)  | - 0.014  | 0.062 | - 0.135                    | 0.108   | 0.986   | - 0.221 | 0.825   |  |
| CS (10€–10€) * direction (pull)   | 0.025    | 0.010 | 0.005                      | 0.045   | 1.025   | 2.398   | 0.016   |  |
| CS (10€–10€) * time (post training)   | 0.008    | 0.010 | - 0.012                    | 0.029   | 1.008   | 0.809   | 0.418   |  |
| Direction (pull) * time (post training)   | - 0.003  | 0.015 | - 0.031                    | 0.026   | 0.997   | - 0.184 | 0.854   |  |
| CS (10€–10€) * training (intervention training)   | - 0.013  | 0.021 | - 0.054                    | 0.028   | 0.987   | - 0.612 | 0.540   |  |
| Direction (pull) * training (intervention training)                                       | 0.020    | 0.030 | - 0.039                    | 0.078   | 1.020   | 0.654   | 0.513   |  |
| Time (post training) * training (intervention training)                                   | - 0.014  | 0.050 | - 0.112                    | 0.085   | 0.987   | - 0.270 | 0.787   |  |
| CS (10€–10€) * direction (pull) * time (post training)                                    | 0.051    | 0.021 | 0.010                      | 0.091   | 1.052   | 2.437   | 0.015   |  |
| CS (10€–10€) * direction (pull) * training (intervention training)                        | 0.012    | 0.021 | - 0.029                    | 0.053   | 1.012   | 0.581   | 0.561   |  |
| CS (10 $(-10\epsilon)$ * time (post training) * training (intervention training)          | - 0.007  | 0.021 | - 0.047                    | 0.034   | 0.993   | - 0.316 | 0.752   |  |
| Direction (pull) * time (post training) * training (intervention training)                | - 0.041  | 0.029 | - 0.099                    | 0.016   | 0.960   | - 1.409 | 0.159   |  |
| CS (10€–10€) * direction (pull) * time (post training) * training (intervention training) | 0.061    | 0.041 | - 0.020                    | 0.142   | 1.063   | 1.474   | 0.140   |  |

**Table 3.** Results of generalized linear mixed model testing the interaction effect between response direction, conditioned stimulus, time, and training on reaction time in the AAT. CS conditioned stimulus, SE standard error, exp(B) odds ratio.



**Figure 3.** There was an overall effect of time, participants responded faster in the post AAT. In addition, after training, there was an increase in RT towards positive CS. *RT* reaction time.

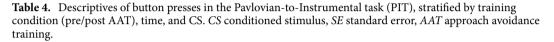
.....

**PIT.** For the pre training PIT, we did find a significant effect of the background CS and instrumental response but no interaction (supplementary Table 4). As expected, subjects responded with more button presses in positive CS trials (mean = 4.30; SE = 0.154) than in neutral CS trials (mean = 3.68; SE = 0.134) and negative CS trials (mean = 3.05; SE = 0.114). In line with this, the effect of instrumental condition was significant as well (*estimate* = 1.192; SE = 0.047; p < 0.001), here button presses in collect trials were higher (mean = 6.19; SE = 0.175) than in not collect trials (mean = 2.14; SE = 0.670).

For summary statistics of PIT pre and post training conditions, please see Table 4.

Scientific Reports | (2023) 13:10074 |

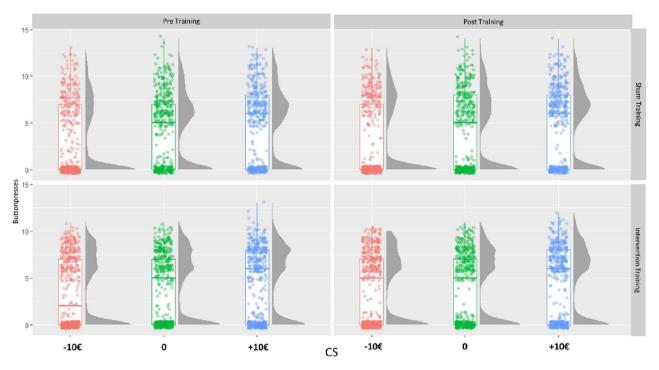
|                       |          |    |       | Button presses |       |  |
|-----------------------|----------|----|-------|----------------|-------|--|
|                       |          |    |       | Mean           | SE    |  |
|                       | Pre AAT  |    | - 10€ | 3.150          | 0.179 |  |
|                       |          | CS | 0     | 4.353          | 0.178 |  |
| Cham turining         |          |    | 10€   | 4.810          | 0.181 |  |
| Sham training         | Post AAT |    | - 10€ | 3.258          | 0.189 |  |
|                       |          | CS | 0     | 4.239          | 0.192 |  |
|                       |          |    | 10€   | 4.906          | 0.186 |  |
|                       | Pre AAT  | CS | - 10€ | 3.613          | 0.165 |  |
|                       |          |    | 0     | 3.984          | 0.164 |  |
| Tutomontion turining  |          |    | 10€   | 4.579          | 0.166 |  |
| Intervention training |          |    | - 10€ | 3.933          | 0.167 |  |
|                       | Post AAT | CS | 0     | 4.032          | 0.166 |  |
|                       |          |    | 10€   | 4.488          | 0.169 |  |



The GLMM showed significant main effects of CS (*estimate* = 0.530; SE = 0.163; p = 0.001) and instrumental condition (*estimate* = 1.30; SE = 0.055; p < 0.001) (*please see* Table 5). The interaction between all fixed effects (CS by training by time by instrumental response) was not significant (*estimate* = 0.330; SE = 0.284; p = 0.247) (please see supplementary Fig. 1), however the significant interaction between CS by training by time (*estimate* = -0.459; SE = 0.144; p = 0.0001) indicates a training-dependent effect of time on the CS-button presses (PIT effect) (*please see* Fig. 4). Simple effects analysis showed a decrease in button presses in negative (*estimate* = -0.362; SE = 0.099; p < 0.001) and neutral CS trials (*estimate* = -0.211; SE = 0.091; p = 0.021) but not positive CS trials (*estimate* = 0.002; SE = 0.089; p = 0.985). This points to an increased PIT effect after sham training. The button presses between training conditions also differed after training, here the response rate in negative CS trials was significantly higher after avoidance compared to sham training (*estimate* = 0.6304; z = 2.048; p = 0.041). In addition, the interaction effect of time by training (*estimate* = 0.448; SE = 0.134; p < 0.001) as well as time by instrumental condition (*estimate* = 0.280; SE = 0.106; p = 0.008) was significant. Here, the overall button presses were significantly higher at pre training compared to post training AAT after sham training (*estimate* = 0.448; SE = 0.104; P = 0.041).

|  |          |       | 95% Exp(B) Confidence<br>Interval |       |         |        |         |
|--|----------|-------|-----------------------------------|-------|---------|--------|---------|
| Names  | Estimate | SE    | Lower                             | Upper | z       | exp(B) | p       |
| (Intercept)  | 0.866    | 0.141 | 1.803                             | 3.135 | 6.137   | 2.377  | < 0.001 |
| Instrumental condition (collect)   | 1.301    | 0.055 | 3.295                             | 4.094 | 23.498  | 3.673  | < 0.001 |
| CS   | 0.530    | 0.163 | 1.236                             | 2.337 | 3.262   | 1.700  | 0.001   |
| Time (post training)   | - 0.138  | 0.067 | 0.764                             | 0.993 | - 2.065 | 0.871  | 0.039   |
| Training (training condition)  | 0.406    | 0.282 | 0.864                             | 2.609 | 1.441   | 1.501  | 0.150   |
| Instrumental condition (collect) * CS  | - 0.129  | 0.076 | 0.758                             | 1.020 | - 1.701 | 0.879  | 0.089   |
| Instrumental condition (collect) * time (post training)                                      | 0.280    | 0.106 | 1.075                             | 1.628 | 2.640   | 1.323  | 0.008   |
| CS * time (post training)  | 0.134    | 0.072 | 0.993                             | 1.317 | 1.858   | 1.143  | 0.063   |
| Instrumental condition (collect) * training (training condition)                             | 0.016    | 0.111 | 0.818                             | 1.261 | 0.142   | 1.016  | 0.887   |
| CS * training (training condition)   | - 0.534  | 0.325 | 0.310                             | 1.108 | - 1.643 | 0.586  | 0.100   |
| Time (post training) * training (training condition)   | 0.448    | 0.134 | 1.205                             | 2.034 | 3.359   | 1.565  | < 0.001 |
| Instrumental condition (collect) * CS * time (post training)                                 | - 0.306  | 0.142 | 0.558                             | 0.973 | - 2.152 | 0.737  | 0.031   |
| Instrumental condition (collect) * CS * training (training condi-<br>tion)                   | - 0.095  | 0.151 | 0.677                             | 1.222 | - 0.630 | 0.909  | 0.529   |
| Instrumental condition (collect) * time (post training) * training (training condition)      | - 0.405  | 0.211 | 0.441                             | 1.008 | - 1.924 | 0.667  | 0.054   |
| CS * time (post training) * training (training condition)                                    | - 0.459  | 0.144 | 0.476                             | 0.838 | - 3.180 | 0.632  | 0.001   |
| Instrumental condition (collect) * CS * time (post training) * training (training condition) | 0.329    | 0.284 | 0.796                             | 2.424 | 1.159   | 1.390  | 0.247   |

**Table 5.** Results of generalized linear mixed model testing the interaction effect of instrumental response, conditioned stimulus, time, and training on button presses in the PIT. CS conditioned stimulus. SE standard error. exp(B) odds ratio.



**Figure 4.** Boxplot/density graph depicting the PIT effect before and after AAT sham/ intervention training. There was a significant interaction effect of CS by time by training, which indicates that the instrumental behavior changed over time due to training. CS = conditioned stimuli.

z=3.360; p<0.001) and button presses in no collect trials were significantly higher before AAT training than after training (*estimate* = 0.280; z=2.641; p=0.008).

There were no effects of training, time, or an interaction on the proportion of correct trials during the forced choice task (please see supplements, supplementary Table 9 and supplementary Fig. 3).

### Discussion

The goal of the current study was the assessment of an AAT modification training and its impact on the PIT effect. We found a significant PIT effect but there was no impact of Pavlovian CS on approach or avoidance bias operationalized with a joystick AAT. Training to avoid positively valenced CS did not result in a significant change in approach and avoidance propensities. In addition, intervention training did not lead to a decreased magnitude of the PIT effect, but an increased response in negative CS trials. On the other hand, participants showed an increased PIT effect after sham training. To our knowledge, this is the first study that suggests a lack of action bias toward experimentally conditioned cues. On the other side, we also show for the first time that training to approach negatively valenced cues reduced the instrumental behavior towards these cues in the PIT transfer phase.

As expected, we found robust significant PIT effects in accordance with previous reports<sup>15,40,49,50</sup>. Instrumental behavior was enhanced by positive Pavlovian cues, while negative cues constrained instrumental responding. Contrary to our hypothesis, the effect of Pavlovian cue presentation did not translate to approach/avoidance behavior, operationalized by joystick RT in the AAT. We expected the positively valued cues to be approached faster and to be avoided slower and vice versa for negatively valued cues. The lack of significant bias towards the conditioned stimuli might be due to various reasons.

For instance, we are using an irrelevant feature version, i.e., participants respond to a feature related to image orientation (horizontal/vertical) and not to the content of the image per se. Studies have made a strong case for the relevant feature version of the AAT for bias measurement<sup>51,52</sup>.

Although effects seem to be maximized by feature-relevant designs, a design in which subjects are not consciously focusing on the stimuli has been proposed to be a more implicit measure of bias<sup>47,53</sup>. However, as a result, our implicit AAT design might not capture possible bias effects as attention is not drawn to the content of the stimuli. In contrast to the transfer part of the PIT task, the attributes of the stimuli per se might not capture the subjects' attention in the AAT. In the PIT task, the Pavlovian cues are tiled over the background while subjects are instructed to acknowledge their presence and their associated value but to respond to the instrumental stimuli only. The AAT instructions refer to the orientation of the image only but not to content—in line with previous AAT training studies<sup>54,55</sup>. In addition, there is no reward for behavior in the AAT while participants receive monetary recompense for both the presentation of the Pavlovian cues as well as instrumental responding (although no direct feedback is given after the trials). Another explanation could relate to the operationalization of the response in PIT and AAT i.e., button pressing versus pushing/pulling of a joystick. Conversely, a PIT paradigm that instrumentalized joystick approach/avoidance movements instead of button presses has been established and robust PIT effects have been found<sup>56</sup>. We also found a significant main effect of time that suggests a learning effect reflected by faster RT in the AAT overall.

Furthermore, instructions are an important predictor of approach bias and overall significance was related to individuals awareness of stimuli valences<sup>53</sup>. From an associative stand, AAT bias was attributed to impulsive, automatic processes and training could alter implicit associations towards cues<sup>23</sup>. This view is now challenged by a vast body of research that propose the idea of inferential processes guiding stimulus-action tendencies<sup>57</sup> According to this theory, these tendencies echo learned instrumental significance that translates into goal-driven behavior<sup>57</sup>. In this context, the importance of instructions can be explained by conscious learning of contingencies of the stimuli and alteration of behavior according to task demands. In line with this, hierarchical theories of the impact of cues on behavior propose that expectancy effects play a large role in driving transfer effects<sup>59</sup>. In their study, Hogarth et al.<sup>59</sup>, observed transfer effects to be attenuated by discriminative extinction training (consisting of extinction of instrumental behavior paired with a CS). However, these transfer effects were also abolished by instructing participants that CS were no indicator of the likelihood of outcome. Here it has also been argued that the efficacy of such interventions is based on "instructed extinction" of instrumental probabilities<sup>13</sup>. The intervention training used in this study uses a similar approach, as the CS condition requires the participants to train a new instrumental avoidance behavior. However, factors such as cognitive instructions or operationalization of behavior (button presses versus joystick movements) alter the context in which the Pavlovian CS exerts possible effects on instrumental behavior<sup>60,61</sup>. This context-specificity could explain the lack of effect of Pavlovian conditioning on the initial pre training AAT performance.

After AAT training the approach and avoidance propensities towards the CS changed—overall we saw that negative CS compared to positive CS were avoided slower and approached faster. However, this was surprisingly not driven by training condition and the sham training group showed the same effect. Again, it is not clear if an approach/avoidance bias towards the CS can be established here, the overall increase in RT indicates a strong learning curve of the task demands and in line with the above, subjects might merely focus on image orientation.

When examining the PIT after training, we did, however, find a significant effect. In this case, contrary to our hypothesis, intervention training did not reduce the PIT effect- however, after sham training, the PIT effect was enhanced. The increased PIT slope after sham training did not fall in line with our expectations. As the transfer phase in this paradigm is done in nominal extinction (i.e., without direct rewarding feedback) and the CS presentation during AAT training was not linked to reward, we did not expect the magnitude of the PIT effect to increase in either condition<sup>62,63</sup>. On the other hand, we found button press response towards negative CS to be increased in the group that underwent intervention training compared to sham training. This might indicate a possible training effect in terms of the increased valence of a previously negatively conditioned cue. Since this is not reflected by RT differences between both conditions during post training AAT, we hypothesize that the AAT might not capture the CS effect on instrumental behavior as the PIT paradigm does. Again, the tasks differ in terms of stimulus presentation, instruction, and operationalization of the instrumental action.

As this is a preliminary proof of concept study, the sample size per condition is rather small and we might have not detected possible effects due to low statistical power. Since PIT effects have been proposed to be larger in clinical populations<sup>17,64</sup>, here the effects of training could have a different impact. However, on the other hand, PIT effects have been consistently found in healthy populations<sup>49,50</sup>. As stated above, any possible changes in instrumental behavior might be due to expectancy effects related to instructions or task set-up, however, we have no qualitative or quantitative information from the participants and think that it would be crucial to implement posthoc questionnaires assessing the participants' subjective judgment on task demands and cognitive strategies. Furthermore, our sample was not balanced in terms of gender distribution, as it contained more women than men. This could potentially affect the generalizability of our results as this imbalance might not accurately reflect prevalence rates in psychiatric conditions, which have been characterized by an altered PIT effect. On the other hand, research has indicated that confounding factors such as gender or age do not influence PIT effects<sup>20</sup>. Our design used monetary reinforcers in instrumental and Pavlovian conditioning, however, different reward types such as e.g., food as a primary reinforcer could elicit stronger or weaker effects on behavior. A study compared the impact of using various reward types and found no difference in PIT effect if their subjective value was matched<sup>65</sup>. In line with this, the monetary rewards might also be valued differently, and future studies should take this into account e.g. by including subjective reward ratings.

In summary, our hypotheses should be investigated in a larger, well-balanced, and clinical sample with i.e., substance abuse or other behavioral patterns marked by conflicts of behavioral goals and values assigned to stimuli.

Taken together we found CS that elicit increased instrumental responding in a PIT task to not affect action bias in an AAT. Sham control training led to a significant increase of the PIT effect. Modification of the PIT effect was seen in that participants that were trained to approach negative CS, showed increased instrumental responses towards those cues in the PIT in contrast to sham training participants. Our findings further contribute to research delineating the underlying mechanism of PIT effects. Some interpretations of our study results refer to the impact of explicit knowledge (i.e., instructions) on how cues energize instrumental behavior. However, further research is necessary to corroborate our findings.

#### Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Received: 1 July 2022; Accepted: 15 June 2023 Published online: 21 June 2023

### References

- 1. Hogarth, L., Dickinson, A. & Duka, T. The associative basis of cue-elicited drug taking in humans. *Psychopharmacology* 208, 337–351. https://doi.org/10.1007/s00213-009-1735-9 (2010).
- Robinson, T. E., Yager, L. M., Cogan, E. S. & Saunders, B. T. On the motivational properties of reward cues: Individual differences. *Neuropharmacology* 76, 450–459. https://doi.org/10.1016/j.neuropharm.2013.05.040 (2014).
- 3. Carter, B. L. & Tiffany, S. T. Meta-analysis of cue-reactivity in addiction research. Addiction 94, 327-340 (1999).
- Jansen, A. A learning model of binge eating: Cue reactivity and cue exposure. Behav. Res. Ther. 36, 257–272. https://doi.org/10. 1016/s0005-7967(98)00055-2 (1998).
- Starcke, K., Antons, S., Trotzke, P. & Brand, M. Cue-reactivity in behavioral addictions: A meta-analysis and methodological considerations. J. Behav. Addict. 7, 227–238. https://doi.org/10.1556/2006.7.2018.39 (2018).
- Heinz, A., Beck, A., Grusser, S. M., Grace, A. A. & Wrase, J. Identifying the neural circuitry of alcohol craving and relapse vulnerability. *Addict. Biol.* 14, 108–118. https://doi.org/10.1111/j.1369-1600.2008.00136.x (2009).
- 7. Heinz, A. et al. Addiction as learned behavior patterns. J. Clin. Med. 8, 1086 (2019).
- Holmes, N. M., Marchand, A. R. & Coutureau, E. Pavlovian to instrumental transfer: A neurobehavioural perspective. *Neurosci. Biobehav. Rev.* 34, 1277–1295. https://doi.org/10.1016/j.neubiorev.2010.03.007 (2010).
- 9. Walker, K. C. The effect of a discriminative stimulus transferred to a previously unassociated response. *J. Exp. Psychol.* **31**, 312–321. https://doi.org/10.1037/h0062929 (1942).
- Geurts, D. E., Huys, Q. J., den Ouden, H. E. & Cools, R. Aversive Pavlovian control of instrumental behavior in humans. J. Cogn. Neurosci. 25, 1428–1441 (2013).
- 11. Garofalo, S. & Robbins, T. W. Triggering avoidance: Dissociable influences of aversive Pavlovian conditioned stimuli on human instrumental behavior. *Front. Behav. Neurosci.* 11, 63 (2017).
- Lewis, A. H., Niznikiewicz, M. A., Delamater, A. R. & Delgado, M. R. Avoidance-based human Pavlovian-to-instrumental transfer. *Eur. J. Neurosci.* 38, 3740–3748 (2013).
- 13. Cartoni, E., Balleine, B. & Baldassarre, G. Appetitive Pavlovian-instrumental transfer: A review. *Neurosci. Biobehav. Rev.* 71, 829–848 (2016).
- 14. Sommer, C. et al. Dysfunctional approach behavior triggered by alcohol-unrelated Pavlovian cues predicts long-term relapse in alcohol dependence. Addict. Biol. 25, e12703. https://doi.org/10.1111/adb.12703 (2020).
- Garbusow, M. et al. Pavlovian-to-instrumental transfer and alcohol consumption in young male social drinkers: Behavioral, neural and polygenic correlates. J. Clin. Med. https://doi.org/10.3390/jcm8081188 (2019).
- Watson, P., Wiers, R. W., Hommel, B., Gerdes, V. E. A. & de Wit, S. Stimulus control over action for food in obese versus healthyweight individuals. *Front. Psychol.* https://doi.org/10.3389/fpsyg.2017.00580 (2017).
- Nord, C. L., Lawson, R. P., Huys, Q. J. M., Pilling, S. & Roiser, J. P. Depression is associated with enhanced aversive Pavlovian control over instrumental behaviour. Sci. Rep. 8, 12582. https://doi.org/10.1038/s41598-018-30828-5 (2018).
- Huys, Q. J. et al. The specificity of Pavlovian regulation is associated with recovery from depression. Psychol. Med. 46, 1027–1035. https://doi.org/10.1017/s0033291715002597 (2016).
- Genauck, A. et al. Cue-induced effects on decision-making distinguish subjects with gambling disorder from healthy controls. Addict. Biol. 25, e12841 (2020).
- Garbusow, M. et al. Pavlovian-to-instrumental transfer across mental disorders: A review. Neuropsychobiology 81, 418–437. https:// doi.org/10.1159/000525579 (2022).
- Mahlberg, J. et al. Human appetitive Pavlovian-to-instrumental transfer: A goal-directed account. Psychol. Res. 85, 449–463. https:// doi.org/10.1007/s00426-019-01266-3 (2021).
- Yee, D. M. & Braver, T. S. Interactions of motivation and cognitive control. *Curr. Opin. Behav. Sci.* 19, 83–90. https://doi.org/10. 1016/j.cobeha.2017.11.009 (2018).
- 23. Wiers, R. W., Gladwin, T. E., Hofmann, W., Salemink, E. & Ridderinkhof, K. R. Cognitive bias modification and cognitive control training in addiction and related psychopathology: Mechanisms, clinical perspectives, and ways forward. *Clin. Psychol. Sci.* 1, 192–212 (2013).
- Cousijn, J., Goudriaan, A. E. & Wiers, R. W. Reaching out towards cannabis: Approach-bias in heavy cannabis users predicts changes in cannabis use (vol 106, pg 1667, 2011). Addiction 106, 2053–2053. https://doi.org/10.1111/j.1360-0443.2011.03677.x (2011).
- Wiers, R., Rinck, M., Dictus, M. & Van den Wildenberg, E. Relatively strong automatic appetitive action-tendencies in male carriers of the OPRM1 G-allele. *Genes Brain Behav.* 8, 101–106 (2009).
- Brockmeyer, T., Hahn, C., Reetz, C., Schmidt, U. & Friederich, H.-C. Approach bias and cue reactivity towards food in people with high versus low levels of food craving. *Appetite* 95, 197–202. https://doi.org/10.1016/j.appet.2015.07.013 (2015).
- Booth, C., Spronk, D., Grol, M. & Fox, E. Uncontrolled eating in adolescents: The role of impulsivity and automatic approach bias for food. *Appetite* 120, 636–643. https://doi.org/10.1016/j.appet.2017.10.024 (2018).
- Eberl, C. et al. Approach bias modification in alcohol dependence: Do clinical effects replicate and for whom does it work best?. Dev. Cogn. Neurosci. 4, 38–51. https://doi.org/10.1016/j.dcn.2012.11.002 (2013).
- Manning, V. et al. Cognitive bias modification training during inpatient alcohol detoxification reduces early relapse: A randomized controlled trial. Alcohol. Clin. Exp. Res. 40, 2011–2019 (2016).
- Rinck, M., Wiers, R. W., Becker, E. S. & Lindenmeyer, J. Relapse prevention in abstinent alcoholics by cognitive bias modification: Clinical effects of combining approach bias modification and attention bias modification. J. Consult. Clin. Psychol. 86, 1005 (2018).
- Boffo, M. et al. Cognitive bias modification for behavior change in alcohol and smoking addiction: Bayesian meta-analysis of individual participant data. Neuropsychol. Rev. 29, 52–78 (2019).
- Kakoschke, N., Kemps, E. & Tiggemann, M. Approach bias modification training and consumption: A review of the literature. Addict. Behav. 64, 21–28. https://doi.org/10.1016/j.addbeh.2016.08.007 (2017).
- Van Gucht, D., Vansteenwegen, D., Van den Bergh, O. & Beckers, T. Conditioned craving cues elicit an automatic approach tendency. *Behav. Res. Ther.* 46, 1160–1169. https://doi.org/10.1016/j.brat.2008.05.010 (2008).
- Thewissen, R., Havermans, R. C., Geschwind, N., van den Hout, M. & Jansen, A. Pavlovian conditioning of an approach bias in low-dependent smokers. *Psychopharmacology* 194, 33–39. https://doi.org/10.1007/s00213-007-0819-7 (2007).
- Olney, J. J., Warlow, S. M., Naffziger, E. E. & Berridge, K. C. Current perspectives on incentive salience and applications to clinical disorders. *Curr. Opin. Behav. Sci.* 22, 59–69. https://doi.org/10.1016/j.cobeha.2018.01.007 (2018).
- Huys, Q. J. M. *et al.* Disentangling the roles of approach, activation and valence in instrumental and Pavlovian responding. *PLoS Comput. Biol.* https://doi.org/10.1371/journal.pcbi.1002028 (2011).
   Watson, P., De Wit, S., Hommel, B. & Wiers, R. Motivational mechanisms and outcome expectancies underlying the approach bias
- Watson, P., De Wit, S., Hommel, B. & Wiers, R. Motivational mechanisms and outcome expectancies underlying the approach bias toward addictive substances. *Front. Psychol.* https://doi.org/10.3389/fpsyg.2012.00440 (2012).
- Chen, K. *et al.* Alcohol approach bias is associated with both behavioral and neural Pavlovian-to-instrumental transfer effects in alcohol-dependent patients. *Biol. Psychiatry Glob. Open Sci.* https://doi.org/10.1016/j.bpsgos.2022.03.014 (2022).
- Chen, K. et al. Automatic approach behaviors in alcohol dependence: Does a cognitive bias modification training affect Pavlovianto-instrumental transfer effects?. Neuropsychobiology 81, 387–402. https://doi.org/10.1159/000526805 (2022).

- Garbusow, M. et al. Pavlovian-to-instrumental transfer in alcohol dependence: A pilot study. Neuropsychobiology 70, 111–121. https://doi.org/10.1159/000363507 (2014).
- 41. Brainard, D. H. & Vision, S. The psychophysics toolbox. Spat. Vis. 10, 433–436 (1997).
- 42. Kleiner, M., Brainard, D. & Pelli, D. What's new in Psychtoolbox-3? (2007).
- 43. Pelli, D. G. & Vision, S. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spat. Vis.* **10**, 437–442 (1997).
- Wiers, R. W., Eberl, C., Rinck, M., Becker, E. S. & Lindenmeyer, J. Retraining automatic action tendencies changes alcoholic patients' approach bias for alcohol and improves treatment outcome. *Psychol. Sci.* 22, 490–497. https://doi.org/10.1177/0956797611400615 (2011).
- 45. Wagenmakers, E.-J. & Farrell, S. AIC model selection using Akaike weights. Psychon. Bull. Rev. 11, 192–196 (2004).
- Payne, E. H., Gebregziabher, M., Hardin, J. W., Ramakrishnan, V. & Egede, L. E. An empirical approach to determine a threshold for assessing overdispersion in Poisson and negative binomial models for count data. *Commun. Stat.-Simul. Comput.* 47, 1722–1738 (2018).
- 47. Zech, H. G., Rotteveel, M., van Dijk, W. W. & van Dillen, L. F. A mobile approach-avoidance task. *Behav. Res. Methods* 52, 2085–2097. https://doi.org/10.3758/s13428-020-01379-3 (2020).
- Zlomuzica, A., Lange, M., Reher, S., Machulska, A. & Rinck, M. The effects of psychological stress on approach tendencies for smoking-related cues in smokers. *Eur. J. Neurosci.* 55, 2581–2591. https://doi.org/10.1111/ejn.15295 (2022).
- Chen, K. et al. The association of non-drug-related Pavlovian-to-instrumental transfer effect in nucleus accumbens with relapse in alcohol dependence: A replication. Biol. Psychiatry 93, 558–565. https://doi.org/10.1016/j.biopsych.2022.09.017 (2023).
- van Timmeren, T. et al. Intact corticostriatal control of goal-directed action in alcohol use disorder: A Pavlovian-to-instrumental transfer and outcome-devaluation study. Sci. Rep. 10, 1–12 (2020).
- Lender, A., Meule, A., Rinck, M., Brockmeyer, T. & Blechert, J. Measurement of food-related approach-avoidance biases: Larger biases when food stimuli are task relevant. *Appetite* 125, 42–47. https://doi.org/10.1016/j.appet.2018.01.032 (2018).
- Kersbergen, I., Woud, M. L. & Field, M. The validity of different measures of automatic alcohol action tendencies. *Psychol. Addict. Behav.* 29, 225–230. https://doi.org/10.1037/adb0000009 (2015).
- Phaf, R. H., Mohr, S. E., Rotteveel, M. & Wicherts, J. M. Approach, avoidance, and affect: A meta-analysis of approach-avoidance tendencies in manual reaction time tasks. *Front. Psychol.* https://doi.org/10.3389/fpsyg.2014.00378 (2014).
- Dickson, H., Kavanagh, D. J. & MacLeod, C. The pulling power of chocolate: Effects of approach–avoidance training on approach bias and consumption. *Appetite* 99, 46–51 (2016).
- 55. Sharbanee, J. M. *et al.* The effect of approach/avoidance training on alcohol consumption is mediated by change in alcohol action tendency. *PLoS ONE* **9**, e85855 (2014).
- 56. Belanger, M. J. *et al.* Development of novel tasks to assess outcome-specific and general Pavlovian-to-instrumental transfer in humans. *Neuropsychobiology* (2022 (accepted)).
- 57. Van Dessel, P., Hughes, S. & De Houwer, J. How do actions influence attitudes? An inferential account of the impact of action performance on stimulus evaluation. *Pers. Soc. Psychol. Rev.* 23, 267–284 (2019).
- Wiers, R. W., Van Dessel, P. & Köpetz, C. ABC training: A new theory-based form of cognitive-bias modification to foster automatization of alternative choices in the treatment of addiction and related disorders. *Curr. Dir. Psychol. Sci.* 29, 499–505 (2020).
- Hogarth, L. et al. Extinction of cue-evoked drug-seeking relies on degrading hierarchical instrumental expectancies. Behav. Res. Ther. 59, 61–70. https://doi.org/10.1016/j.brat.2014.06.001 (2014).
- León, S. P., Abad, M. J. & Rosas, J. M. Giving contexts informative value makes information context-specific. *Exp. Psychol.* 57, 46–53 (2009).
- Rosas, J. M., Todd, T. P. & Bouton, M. E. Context change and associative learning. Wiley Interdiscip. Rev. Cogn. Sci. 4, 237–244. https://doi.org/10.1002/wcs.1225 (2013).
- Lovibond, P. F., Satkunarajah, M. & Colagiuri, B. Extinction can reduce the impact of reward cues on reward-seeking behavior. Behav. Ther. 46, 432–438. https://doi.org/10.1016/j.beth.2015.03.005 (2015).
- 63. Bezzina, L., Lee, J. C., Lovibond, P. F. & Colagiuri, B. Extinction and renewal of cue-elicited reward-seeking. *Behav. Res. Ther.* 87, 162–169. https://doi.org/10.1016/j.brat.2016.09.009 (2016).
- Garbusow, M. et al. Pavlovian-to-instrumental transfer effects in the nucleus accumbens relate to relapse in alcohol dependence. Addict. Biol. 21, 719–731. https://doi.org/10.1111/adb.12243 (2016).
- Lehner, R., Balsters, J. H., Herger, A., Hare, T. A. & Wenderoth, N. Monetary, food, and social rewards induce similar Pavlovianto-instrumental transfer effects. Front. Behav. Neurosci. https://doi.org/10.3389/fnbeh.2016.00247 (2017).

# Author contributions

A.B. and N.R.-S. conceptualized the study. A.R. wrote the first draft of the manuscript and prepared all tables and figures. A.B., N.R.-S., K.C. and A.R. reviewed and edited the manuscript. A.R. finalized and submitted the manuscript.

# Funding

Open Access funding enabled and organized by Projekt DEAL. The funding was provided by CSC Grant 201806750014. This work was partially supported by the German Research Foundation (DFG, Project-ID 402170461–TRR 265; 390688087–EXC 2049).

# **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/s41598-023-37083-3.

Correspondence and requests for materials should be addressed to A.R.

Reprints and permissions information is available at 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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2023