# Devaluation of response-produced safety signals reveals circuits for goal-directed versus habitual avoidance in dorsal striatum.

Robert M. Sears[1,2,4], Erika Andrade[2,4], Shanna B. Samels[2], Lindsay C. Laughlin[2], Danielle M. Moloney[1], Donald A. Wilson[1,2], Matthew R. Alwood[3], Justin M. Moscarello[3] and Christopher K. Cain[1,2]

[1]Department of Child & Adolescent Psychiatry

NYU Grossman School of Medicine

1 Park Avenue, 8th Floor, New York, NY 10016

[2]Emotional Brain Institute

Nathan Kline Institute for Psychiatric Research

140 Old Orangeburg Road, Orangeburg, NY 10962

[3]Department of Psychological & Brain Sciences

Texas A&M Institute for Neuroscience

Texas A&M University

301 Old Main

TAMU MS 3474

College Station, TX 77843-3474

[4]These authors contributed equally

Corresponding author:

Christopher Cain, PhD

Christopher.Cain@NKI.RFMH.org

(845) 398-6560

## Abstract

Active avoidance responses (ARs) are instrumental behaviors that prevent harm. Adaptive ARs may contribute to active coping, whereas maladaptive avoidance habits are implicated in anxiety and obsessive-compulsive disorders. The AR learning mechanism has remained elusive for almost a century, as successful avoidance trials produce no obvious reinforcer. We used a novel outcome-devaluation procedure in rats to show that ARs are positively reinforced by response-produced feedback (FB) cues that develop into safety signals during training. ARs in males were sensitive to FB-devaluation after moderate training, but not overtraining, consistent with a transition from goal-directed to habitual avoidance. Using chemogenetics and FB-devaluation, we also show that goal-directed vs. habitual ARs depend on dorsomedial vs. dorsolateral striatum, suggesting a significant overlap between the mechanisms of avoidance and rewarded instrumental behavior. Females were insensitive to FB-devaluation due to a remarkable context-dependence of counterconditioning. However, degrading the AR-FB contingency suggests that both sexes rely on safety signals to perform goal-directed ARs.

Active avoidance (AA) is a form of response learning that allows subjects to suppress fear and neutralize threats. AA likely evolved to expand the defensive behavior repertoire beyond a narrow set of innate, inflexible reactions (i.e. fight, flight, freezing). Avoidance response (AR) learning may even underlie complex defensive behaviors to wide-ranging threats that could not have evolved under predatory pressure. For instance, a commuter may buckle their seatbelt to protect against future injury or slam their brakes to prevent imminent harm. Clinically, AA is important because adaptive ARs may contribute to active coping and resilience[1,2]. Conversely, inappropriate or persistent avoidance habits are thought to protect irrational fears from extinction[3] and contribute to disorders of control like OCD[4].

AA is typically studied by presenting subjects with a warning stimulus (WS; e.g., sound) that predicts the delivery of an aversive unconditioned stimulus (US; e.g., shock). This transforms the WS into a conditioned threat that arouses fear and triggers reactions like freezing. Subjects eventually emit the AR (e.g., barpress) and gain exposure to avoidance contingencies. As training continues, fear reactions to the WS are progressively replaced by ARs.

AA was popular for studying psychological mechanisms of learning in the mid-20$^{th}$ century. However, progress in understanding the associative structures and neurobiological mechanisms of AA was stalled by theoretical disputes about the reinforcement mechanism and misconceptions about the mediating role of fear[5,6]. Prominent theories focused on stimulus-response (S-R) learning via negative reinforcement (i.e. fear-reduction at WS-termination vs. US-omission) [7,8]. AA research also failed to fully capitalize on modern conceptions of instrumental behavior developed using appetitive reinforcers[9,10]. This work moved beyond the notion that instrumental training produces only reflexive S-R habits. Subjects also learn goal-directed responses that depend on the action-outcome (A-O) contingency and outcome value. Goal-directed responding is typically diagnosed with outcome-devaluation or contingency degradation procedures[9,10]. S-R habits are usually inferred by insensitivity to these tests. However, the situation is more complicated for AA as there is little agreement about what reinforces ARs. If WS-termination and/or US-omission are the key outcome(s), it is unclear how one could verifiably devalue these outcomes, or deliver more of them. Identifying a clear reinforcement mechanism would establish that AA is instrumental and help differentiate between goal-directed versus habitual ARs.

An interesting and less studied hypothesis is that ARs are positively reinforced by response-produced safety signals that are generated via their inverse relationship with the WS and US[8,11-13]. Explicit feedback (FB) cues that immediately follow ARs develop conditioned inhibition (safety) and reinforcing properties[11,14-17]. It is unclear, however, if safety signals behave as valued outcomes supporting goal-directed avoidance, akin to food delivery in a barpressing experiment. To our knowledge, there have been no AA studies that have attempted to devalue response-produced safety signals. To test this, we designed a novel devaluation protocol where AR-FB cues are counterconditioned by pairing with shock. We also present data on a novel contingency degradation test, where non-contingent FB cues are presented during avoidance.

If ARs are positively reinforced by safety, this raises the exciting possibility that AA relies on instrumental brain circuits already identified by appetitive studies. Moderately trained instrumental responses are sensitive to outcome devaluation and are suppressed by manipulations that impair neural activity in posterior dorsomedial striatum (pDMS) [18,19]. In contrast, instrumental responses that are overtrained or trained on interval schedules are insensitive to outcome devaluation and are suppressed by manipulations that impair dorsolateral striatum (DLS) [19,20]. Importantly, however, this latter effect is only evident in subjects that also received outcome devaluation or contingency degradation. This is because goal-directed responding mediated by a parallel pDMS circuit regains

control when DLS-mediated habits are impaired[20]. Thus, if goal-directed and habitual avoidance memories co-exist in different brain circuits, a tool like outcome devaluation may be necessary to identify and study these circuits. Impairing either circuit alone may produce no change in overt avoidance behavior.

Here, we show that counterconditioning of response-produced safety signals impairs moderately trained, but not overtrained, ARs in male rats. We also demonstrate that moderately trained ARs depend on pDMS, and overtrained ARs depend on DLS, as in appetitive studies. Further, we found that ARs in female rats are insensitive to devaluation due to a remarkable context dependence of counterconditioning. Lastly, we show contingency degradation data suggesting that both sexes rely on safety signals for goal-directed avoidance.

## Results

**Avoidance is acquired similarly in male and female rats.** Overtraining is frequently used to induce habitual responding in studies where responses are rewarded[21]. To study the development of avoidance habits, we used a signaled active avoidance (SigAA) shuttlebox procedure in rats and varied the amount of training. Fig. 1 shows combined SigAA training data from multiple behavioral experiments. The percentage of good avoiders was similar in male and female rats (Fig. 1b). Moderately trained good avoiders approached 100% successful avoidance by the fifth training session and average AR latencies settled around 5s after WS onset (Fig. 1c). This was maintained during overtraining (Fig. 1d). Interestingly, variation in AR latencies declined significantly with overtraining even as average AR latencies remained constant (Fig. 1d, right). This relatively invariant response topography is a positive sign of automaticity[22], and possibly habitual control[23]. Nearly identical patterns of AR learning and performance were observed for males and females in both the moderately trained and overtrained cohorts. The only sex differences observed were slightly more shuttles during the acclimation and inter-trial intervals (ITIs) for female rats, suggesting that females are more active than males (Extended Fig. 1).

**Devaluation of response-produced safety signals impairs moderately trained avoidance in male rats.** To test whether goal-directed avoidance depends on safety signals, we devised a novel devaluation protocol where AR-FB tones are counterconditioned with shock. Male rats received moderate SigAA training followed by paired (FB-Devalued and FB-Control groups) or explicitly unpaired (FB-Valued group) counterconditioning in a separate chamber (Fig. 2a). The unpaired procedure controlled for stimulus exposure during counterconditioning but preserved the safety value of the FB tone. The FB-Control group was treated identically to FB-Devalued group except the FB stimulus used during SigAA training was different (house light off). This allowed us to evaluate whether FB tones are transformed into safety signals during SigAA training, by assessing *retardation of acquisition*[24] during counterconditioning compared to rats that also had SigAA training, but the tone was novel. It also allowed us to assess whether any AR impairment induced by counterconditioning was specific to the FB stimulus used during SigAA training. After counterconditioning, rats were returned to the shuttleboxes for two devaluation tests (15 presentations of the WS). Shuttling had no effect on WS duration, and no FB or shocks were presented. After this, a final Tone Fear Test was conducted in a novel open field to assess whether counterconditioning was effective (FB-Valued vs. FB-Devalued) and to assess retardation of acquisition (FB-Devalued vs. FB-Control) in a neutral context.

At the end of SigAA training, rats in all moderately trained groups displayed similarly high levels of avoidance responding (Fig. 2b). Freezing measured during the first five trials of counterconditioning revealed several important group differences (Fig. 2c). First, tone-freezing increased quickly in the FB-Devalued group but remained low in the FB-Valued group, indicating that paired counterconditioning had the intended effect. Second, tone-freezing in the FB-Control group on Trial 2 was higher than FB-Devalued rats, consistent with retardation of acquisition for FB-Devalued rats that began the session treating the tone as a safety signal. Third, freezing in the FB-Valued group was high before tone presentations, due to context fear induced by unpaired shocks, but ceased during tone presentations (Fig. 2c, right). This dramatic *summation* effect[24] is another indicator that tones used as FB during SigAA training become strong safety signals. These conclusions were supported by freezing behavior during the Tone Fear Test in the novel open field (Fig. 2d). FB-Devalued rats showed tone-freezing intermediate to high-freezing FB-Control rats and low-freezing FB-Valued rats. Behaviors during the Devaluation Tests indicate that FB counterconditioning significantly impaired avoidance after moderate training, especially during the first testing block, when responding is most clearly related to a memory of the instrumental outcome[25], before the potential influence of new learning. Shuttling during the WS was much lower for the FB-Devalued group compared to both control groups (Fig. 2e). Consistent with this, AR latencies were significantly longer for FB-Devalued rats (Fig. 2f). Finally, we also measured freezing to the WS during the first block of testing. Freezing is a fear-related behavior that competes with anxiety-related avoidance[5], and multiple studies have found that suppressing AA disinhibits freezing[26-28]. This was also true in our experiment as FB-Devalued rats froze significantly more than both control groups during WS presentations. Together, these data indicate that avoidance FB stimuli become safety signals with SigAA training, these safety signals are devalued by counterconditioning, and this devaluation impairs moderately trained avoidance responding even when the devalued FB stimulus is not presented during the test. This is consistent with goal-directed responding after moderate SigAA training where the decision to avoid depends on a memory of the current outcome value. Importantly, data from the FB-Control group also show that our devaluation effect was specific; strong tone-shock conditioning had no effect on ARs trained with a different FB stimulus.

We next conducted the same analyses in overtrained male rats subjected to the same counterconditioning procedure. Two additional Tone Fear Tests were conducted in the shuttleboxes and counterconditioning boxes to address sex differences (discussed later). Rats assigned to FB-Devalued and FB-Valued groups displayed similarly high levels of avoidance during the 20th SigAA training session (Fig. 2h). Freezing patterns early in the counterconditioning session closely resembled those observed after moderate training. Tone-freezing rose quickly in the FB-Devalued group but remained low in the FB-Valued group (Fig. 2i, left). Rats in the FB-Valued group also showed strong suppression of context freezing during tone presentations (summation effect; Fig. 2i, right). Tone Fear Tests confirmed that counterconditioning was effective and transferred to all contexts tested (Fig. 2j). However, unlike moderately trained rats, FB counterconditioning had no effect on overtrained avoidance during the Devaluation Tests. Shuttling during the WS was equally high for both groups (Fig. 2k, left), including during the critical first block (Fig. 2k, right). Consistent with this, AR latencies (Fig. 2l) and WS-freezing (Fig. 2m) were also unaffected. These data indicate that FB counterconditioning remains effective after overtraining, but ARs are insensitive to devaluation, most likely because they are mediated by S-R associations that bypass outcome value when controlling behavior.

**pDMS is required for goal-directed active avoidance.** To test the role of pDMS, aDMS, and DLS in moderately trained SigAA we used a viral κ opioid receptor DREADD (KORD) strategy to reversibly

suppress neuronal activity in dorsal striatum. KORD is an engineered Gαi-coupled receptor with no natural ligand that hyperpolarizes neurons when bound to Salvinorin B (SalB), an otherwise inert compound[29,[1]]. We first expressed KORD unilaterally in dorsal striatum neurons (see Methods). We then monitored field activity in awake behaving rats via bilateral chronically implanted electrodes before and after SalB (i.v., 2.5 mg/kg; Extended Fig. 3a). Theta band oscillations are a primary LFP signature of activity in dorsal striatum[30-32], and neurons in dorsal striatum become entrained to theta-band frequencies during instrumental learning[30]. Based on assessment of the full mean FFT spectrum, the largest effect of SalB was suppression of theta (9-15Hz) in the KORD hemisphere (Extended Fig. 3c). Spontaneous theta was significantly depressed compared to baseline and compared to the control hemisphere (Extended Fig. 3d). Theta in the control hemisphere was not significantly affected by SalB. Together, these data confirm that KORD/SalB selectively suppresses behaviorally relevant neural activity in dorsal striatum.

We next expressed KORD bilaterally in pDMS of male rats before moderate SigAA training (Fig. 3a,b). Each rat then received two AR test sessions after Vehicle or SalB injections in a counterbalanced order. AR tests consisted of 15 presentations of the WS where shuttling had no effect on WS duration, and no FB stimuli or shocks were presented. At the end of SigAA training, rats assigned to receive Vehicle-first or SalB-first displayed equivalent avoidance (Fig. 3c). Compared to Vehicle, SalB significantly reduced shuttling during the WS, especially during the first testing block (Fig. 3d). There was also a trend towards longer AR latencies during the SalB test (Block 1: $p=0.11$, Fig. 3e). Consistent with an impairment of avoidance, freezing to the WS during block 1 of testing was increased by SalB (Fig. 3f). In contrast, SalB had no effect on moderately trained avoidance or WS-freezing in rats expressing KORD in aDMS (Extended Fig. 3f-j) or DLS (Fig. 3g-k). Further, SalB also had no effect in overtrained rats expressing KORD bilaterally in pDMS (Extended Fig. 3k-o), presumably because S-R associations stored outside pDMS maintain responding. This aligns with studies showing that pDMS lesions fail to impair habitual responding trained with food rewards[20]. Together, these findings indicate that pDMS is critical for performance of goal-directed, but not habitual, ARs. They also suggest that DMS-mediated control of goal-directed ARs occurs before the development of DLS-mediated habits (Fig. 3l-n). Note that KORD/SalB occasionally reduced shuttles during the acclimation or ITI periods of extinction tests, suggesting a modest influence on overall activity (Extended Figs. 3,4). However, these effects were inconsistent and can't explain the patterns of responding observed during the WS.

**DLS is required for habitual active avoidance.** Many studies using appetitive reinforcers have implicated DLS in instrumental habits[33,34]. DLS lesions have been tested in SigAA, but devaluation tests were not used and results were inconclusive[35]. According to dual process accounts of instrumental behavior, goal-directed and habitual associations co-exist after overtraining and are expressed via parallel circuits that include pDMS and DLS[36]. Thus, impairing DLS activity may not reduce overt avoidance behavior, as this will simply disinhibit goal-directed responding (Fig. 4l,m). To implicate DLS in habitual responding, a tool like outcome devaluation or contingency degradation is necessary to also undermine goal-directed motivation and produce a decrement in overt responding (Fig. 4o).

To test whether DLS contributes to avoidance habits, we expressed KORD bilaterally in DLS, overtrained all rats, and used paired counterconditioning to devalue FB tones in half the rats (FB-Devalued group). The remaining rats received unpaired counterconditioning (FB-Valued group). As above, rats then received two Tests consisting of 15 WS presentations after counterbalanced Vehicle or SalB injections. At the end of SigAA overtraining, rats assigned to receive Vehicle- or SalB-first displayed equally high levels of avoidance responding (Fig. 4c,h). In the FB-Valued group, SalB had no

effect on WS-shuttles, AR latencies, or WS-freezing (Fig. 4d-f). However, in the FB-Devalued group, SalB impaired WS-shuttles, increased AR latencies, and increased WS-freezing (Fig. 4i-k). This pattern of results supports the hypothesis that parallel circuits in dorsal striatum control goal-directed vs. habitual ARs, with DLS being critical for habitual ARs (Fig. 4l-o).

**Devaluation of response-produced safety signals fails to impair avoidance in female rats.** Our initial behavioral experiments used groups that were 50% male and 50% female. However, these were underpowered to detect sex differences and females seemed less sensitive to FB devaluation. We then added rats to all groups until we had fully powered analyses of devaluation after moderate training vs. overtraining in both sexes. Male data are reported in Fig. 3 and Extended Fig. 3, showing a sensitivity to FB devaluation after moderate training, but not overtraining. An identical analysis in females shows insensitivity to FB devaluation, regardless of the amount of training. Females assigned to FB-Devalued or FB-Valued groups showed equally high avoidance rates at the end of SigAA training (Fig. 5b,h). During counterconditioning, tone-freezing rose quickly in FB-Devalued rats but remained low in FB-Valued rats (Fig. 5c,i), similar to male rats (Fig. 2c,i). During the Devaluation tests, WS-shuttles (Fig. 5e,k), AR latencies (Fig. 5f,l), and WS-freezing (Fig. 5g,m) were similar in both groups, consistent with unimpaired avoidance behavior even after moderate training.

**Unlike males, counterconditioning is highly context-dependent in female rats.** Note that tone-freezing data during counterconditioning strongly suggests that our FB devaluation protocol was as effective in females as males (compare Fig. 5c,i to Fig. 2c,i). Thus, there are at least three possible explanations for the failure to impair moderately trained avoidance with FB devaluation in females: 1) females transition to habits sooner, 2) females rely on a different outcome for goal-directed avoidance (e.g., US omission), or 3) counterconditioning of safety signals does not transfer to other contexts in females. Tone Fear Tests conducted after Devaluation Tests strongly support the third explanation. Moderately trained females subjected to FB devaluation showed very low tone-freezing in the novel open field (14%; Fig. 5d). To explore this further, we added additional Tone Fear Tests in the shuttleboxes and the counterconditioning boxes for overtrained rats. Although FB-Devalued females showed more tone-freezing than FB-Valued females in all locations (Fig. 5j), average freezing was extremely low for females tested in the open field (FB-Devalued: 9%, FB-Valued: 3%) and shuttleboxes (FB-Devalued: 5%, FB-Valued: 2%). The only place where females showed substantial tone-freezing after FB devaluation was in the counterconditioning boxes (53%; Fig. 5j, right). Thus, unlike males (Fig. 2d,j), females only express counterconditioning of safety signals in the context where counterconditioning occurred.

To circumvent the context-dependence of counterconditioning in females, we conducted three additional experiments where females were moderately trained, and counterconditioning occurred in the shuttleboxes. In the first experiment, counterconditioning was identical to previous experiments except it occurred in the shuttleboxes with the door open (Extended Fig. 5a-f). This appeared to solve the context-dependence problem, as Tone Fear Tests showed high tone-freezing for FB-Devalued rats in the shuttleboxes even though tone-freezing in the open field remained very low. However, avoidance behavior seemed to collapse during the Devaluation Tests and there were no differences between groups on any measure. The second experiment was identical to the first except we used a weaker counterconditioning procedure (20 tones and shocks, paired or unpaired) to maintain avoidance responding for FB-Valued rats (Extended Fig. 5g-l). This produced a nearly identical pattern of results.

In the third experiment, we again used the weaker counterconditioning procedure, but closed the shuttlebox door during counterconditioning (Extended Fig. 5m-r). This was done because we noticed some rats attempting to shuttle during the counterconditioned tone, and it is unclear how this might affect WS-shuttling at test. Although this led to higher levels of avoidance during the Devaluation Tests, we again observed no differences between FB-Devalued and FB-Valued females on any measure. Interestingly, FB-Devalued females showed high tone-freezing in the shuttleboxes, but only when the door was closed – as during counterconditioning (Extended Fig. 5r). Together, these data indicate a remarkable context-dependence of safety signal counterconditioning in female rats.

**Degrading AR-safety contingency impairs moderately trained avoidance in both sexes.** In appetitive studies, goal-directed responding is also sensitive to A-O contingency degradation (CD). For instance, rats barpressing for food will reduce responding if food is delivered for free (non-contingently), even if food remains available upon pressing[37]. CD may be ideal for assessing goal-directedness of ARs in females because it preserves the value of the outcome (FB=safety) and occurs in the training context. We designed a CD protocol using male rats shuttling in an unsignaled active avoidance (USAA) paradigm[27]. During USAA, there is no explicit WS and shocks are scheduled for delivery every 5s. Each shuttle produces a FB cue and delays footshock by 30s. Shuttles during response-shock intervals are coded as ARs and shuttles following a shock are coded as escape responses (ERs). Explicit response-produced FB cues improve learning[38] and become safety signals with USAA training[39]. After reaching training criterion, subjects were matched into equal Control (CTL) and CD groups. CTL rats received four additional USAA training sessions. CD rats received similar sessions, except free safety signals (+30s of safety from shock) were delivered non-contingently (see Methods and Fig. 6a). This reduced ARs and ERs by about 50% (Fig. 6b,c) while having no significant effect on the frequency of shocks (Fig. 6d). By design, CD rats also received significantly more safety signals than CTL rats (Fig. 6e). These data converge with our FB devaluation data (Fig. 2), further implicating safety signals in goal-directed avoidance.

To test whether females are also sensitive to CD after moderate USAA training, parallel experiments were conducted at two institutes by different experimenters using identical protocols and equipment. Results are combined in Fig. 6f-g and shown separately in Extended Fig. 6. Like males, CD females received more safety signals and emitted significantly fewer ARs and ERs (Fig. 6i,f,g). CD females also received fewer shocks than males (Fig. 6h), likely because they shuttle at a higher rate throughout the sessions. Remarkably similar patterns of ARs, ERs, shocks, and safety signals were observed at both institutes (Extended Fig. 6), indicating that this CD effect is robust. These data suggest that females also rely on FB safety signals for goal-directed avoidance after moderate training.

## Discussion

We found that response-produced FB stimuli are transformed into safety signals during AA training that support goal-directed avoidance. They also show that overtrained ARs are insensitive to FB devaluation. Together, these data confirm that AA can indeed produce instrumental responses directed towards safety, a topic that has been long-debated[5,13,40,41] but, until now, never demonstrated empirically. These findings, along with our KORD/SalB studies, are also consistent with associative dual process accounts of instrumental behavior implicating parallel dorsal striatum circuits in goal-directed vs. habitual instrumental behavior[10,33,42]. This pattern of results suggests that AA relies on

similar psychological and neural mechanisms as other positively reinforced instrumental behaviors once FB stimuli become valued safety signals. Our data also confirm that freezing and ARs are anticorrelated, suggesting that AA is an anxiety-related response to threat that is distinct from fear responses and selected in contexts where safety is attainable. These results challenge traditional theories that frame avoidance behavior as reflexive (S-R), negatively reinforced, and fear motivated.

A core finding in appetitive instrumental research is that positively reinforced behavior controlled by goal-directed vs. habitual associations depends on parallel, competitive circuits in dorsal striatum. pDMS is required for devaluation-sensitive responses and DLS is required for devaluation-insensitive responses[18,20,36,43]. Little is known about the circuits mediating AA, however, several reports are consistent with the idea that AA is reinforced by safety and controlled by different brain circuits after overtraining. AA learning and performance also depends on lateral and basal amygdala[27,44-46], and on the serial connection between basal amygdala and nucleus accumbens shell (NAcS) [26]. However, overtrained ARs are amygdala-independent[27,45]. Avoidance behavior can also be bidirectionally modulated by manipulating dopamine prediction error signals to AR-FB in nucleus accumbens after moderate training, but not overtraining[47-49]. NAcS has also been implicated in the positive reinforcement of overtrained ARs by safety[50]. Finally, Darvas and colleagues showed that replacing dopamine in amygdala+striatum rescued the ability of dopamine-deficient mice to acquire ARs[51]. However, once learned, dopamine was required only in striatum to perform the response.

Studies of pDMS in goal-directed avoidance are sparse and inconclusive. For instance, one study using pretraining lesions found a modest impairment in the rate of AR learning[35]. However, it is possible that compensation by other regions like DLS enabled learning. Operating on the assumption that goal-directed associations are acquired more rapidly than habitual associations[19], we tested the hypothesis that pDMS is required for avoidance after moderate training – a phase shown earlier to be sensitive to FB-devaluation (Fig. 2). We found that KORD-mediated inhibition of pDMS reduced ARs during the WS and caused a return of WS-induced freezing after moderate training. Interestingly, we found no effect on AR performance when aDMS or DLS was suppressed after moderate training, or when pDMS was suppressed after overtraining. This anatomically-specific and training-sensitive effect provides strong support that goal-directed avoidance depends on the same posterior region of DMS that mediates appetitively reinforced instrumental actions[18]. Also consistent with appetitive studies, inhibition of DLS impaired avoidance only in rats that also received FB-Devaluation. This suggests that avoidance relies on the same parallel circuits in dorsal striatum that mediate instrumental actions vs. habits. It also demonstrates the utility of our FB devaluation paradigm for identifying goal-directed behaviors and isolating avoidance habit circuits for study. Without a tool like FB devaluation, manipulations of DLS or other parts of the habit circuit may have no effect on overt avoidance.

Although negative reinforcement theories dominated during the heyday of AA research, the idea that safety signals may positively reinforce ARs was introduced as early as 1936 by Konorski and Miller. This was echoed by Soltysik in 1963, who called inhibitory AR-FB a “secondary reward” [12] that “stabilizes avoidance behavior and makes it virtually independent of the noxious US” [52]. Studies since then provide indirect support for the notion that safety signals reinforce ARs without directly demonstrating an AR-FB association[11,16,17,38,53-55]. Perhaps most convincingly, Weisman and Litner trained rats separately on USAA and Pavlovian safety conditioning[14]. They next modified the response requirement to prevent shock and introduced the safety cues as AR-FB. FB increased responding when the requirement was high and decreased responding when the requirement was low. Though these studies are consistent with the idea that inhibitory FB cues positively reinforce ARs, no distinction was made between goal-directed vs. habitual ARs. Since AR-FB cues are presented during the critical tests,

it is impossible to know if subjects are responding based on knowledge of an action-outcome relation and memory for the outcome value. It also remains unclear if effects like those observed by Weisman and Litner would disappear with overtraining.

Surprisingly few studies have attempted to distinguish between goal-directed and habitual AA. Initial work in this vein focused on devaluing aversive USs after avoidance training[4,56-59]. Devaluation strategies included reducing the chamber temperature for rats avoiding heat[56], disconnecting shock wires for humans avoiding shock[4,57,58], or pairing shocks with morphine[59]. Although decrements in avoidance responding were observed, these results are difficult to interpret. None show that AR decrements occur from the beginning of devaluations tests, when responding is most clearly linked to a memory of the outcome value[25,60]. Amount of training was typically not varied and none show a loss of devaluation effects with overtraining. Rodent studies did not verify that US value was reduced outside the avoidance situation. Although these data suggest that current shock value contributes to avoidance, one study's authors concede that the precise way this happens "remains undetermined" [59]. Finally, it is possible that devaluing an aversive US affects ARs indirectly by reducing the value of safety cues that predict its absence. Indeed, some have described the conditioned inhibition that defines safety signals as a "slave" process to aversive excitation[61].

We directly tested the safety signal hypothesis of goal-directed avoidance by creating a novel outcome-devaluation procedure. Our study design addressed the above issues and was inspired by popular devaluation protocols developed for appetitive studies[19]. The near complete collapse of avoidance after devaluation suggests that response-produced safety signals are the primary outcome supporting AR performance after moderate training. These results may help to resolve problems that slowed progress in understanding AA. The finding that safety signals reinforce avoidance may explain how WS-termination and US-omission contribute to avoidance behavior[62,63]. These events are inversely related to AR-FB cues and likely both support Pavlovian safety learning[2]. They also help explain the extreme resistance to extinction sometimes observed after avoidance training[64]. Discontinuing the US does nothing to undermine AR→FB or WS→AR associations that should maintain responding. These results may also clarify why some ARs are learned readily while others are not (e.g. barpress avoidance). Responses that produce the greatest stimulus change likely produce the strongest safety signals[65]. Our results are also consistent with many studies observing that measures of fear are poorly correlated with avoidance[5,7]. As ARs are successfully acquired, the meaning of the WS likely changes from a threat of imminent harm to a threat of harm in a context where safety is easily attainable[5,66]. This should induce a shift from a central state of fear to one of anxiety, where harm is less likely and volitional action is possible[67,68]. Since fear and anxiety are distinct defensive brain states[69], innate fear reactions like freezing are suppressed or remain untriggered during anxiety states that support avoidance. However, if an effective AR is blocked or otherwise impaired, the fear state returns along with inflexible reactions like freezing[70] (see also Figs. 2,3,4).

Although avoidance behavior was very similar in males and females, we found a major sex difference in our FB-Devaluation studies. Females showed no decrement in avoidance responding when the FB-tone was counterconditioned after moderate training. This was not due to a failure of safety conditioning or a failure of counterconditioning. Instead, females displayed a remarkable degree of context dependence in their expression of counterconditioning that was not observed in males. Most surprising, when we attempted to conduct FB-Devaluation in the shuttleboxes to circumvent this context dependence, females behaved as if the shuttlebox with the door open vs. closed were separate contexts. Tone-freezing for FB-Devalued rats was only 13% with the door open but 68% with the door

closed. This apparent all-or-none contextual gating of counterconditioning means that females treat AR-FB as safe everywhere but the exact location where counterconditioning occurred. We cannot find any previous reports of this sex difference in the literature. However, studies on a different form of counterconditioning were conducted primarily in females and the expression of counterconditioning was context-dependent[71]. This effect appears to be specific to counterconditioning. Other forms of renewal, like the return of freezing after extinction, seem slightly weaker in females[72,73]. As a practical matter, our result suggests that FB-Devaluation may not be a useful tool for distinguishing goal-directed and habitual ARs in female rats.

In our final experiment, we used a novel contingency degradation protocol after USAA training to confirm our behavioral findings in males and determine if females also rely on safety signals for goal-directed avoidance. After moderate training, males adjusted their AR rates downward when free (non-contingent) FB safety signals were presented during training. This was replicated in females by different experimenters at different institutions, suggesting the effect is robust. Thus, contingency degradation protocols may be especially useful for studying neural circuits of goal-directed vs. habitual avoidance in females. However, more work is needed to determine if delivering free safety signals loses effectiveness with overtraining.

In summary, this work suggests that volitional action is indeed possible under threat, even intense threat[74,75], provided that responses are available in the context to attain safety[67]. Our safety-signal devaluation protocol provides a new tool for differentiating between goal-directed and habitual ARs in neurobiological studies and in studies of aberrant avoidance habits in humans. Recent work using US-devaluation protocols suggests that patients with OCD[4], or with a history of early life stress[58], have slightly stronger avoidance habits. It would be interesting to see if evidence for more robust AR habits is obtained after safety signal devaluation. Our data showing strong context-dependence of counterconditioning in females may also have clinical implications if replicated in humans. Experiences that devalue an avoidance reinforcer may only reduce avoidance in the contexts where they happen, possibly contributing to the higher rate of anxiety disorders characterized by avoidance in females. Interestingly, a recent human avoidance study found that ratings of “relief pleasantness” during response-produced safety periods relate to reward prediction error signals that decline in most individuals as the AR is acquired[76]. However, individuals scoring low on distress tolerance showed no such decline, suggesting possible over-reinforcement of avoidance habits. A safety signal devaluation test may help resolve this question and potentially identify a new mechanism contributing to anxiety. Lastly, our results provide empirical support for reinforcement learning models that rely on response-produced safety signals to explain AA[77].

## Footnotes

[1] Although SalB is reported to be inert, we noticed effects of SalB injections in some of our male rats, including naïve rats not expressing KORD. Effects included a suppression of locomotor activity, squinting, and labored breathing that did not occur with vehicle injections. The effects appeared to subside within 10-20 minutes and could not account for the avoidance effects we observed with SalB+KORD. We did not observe similar effects in females who all received less total drug than males.

[2] Our results do not rule out additional roles for WS-termination and/or US-omission in the negative reinforcement of AA, especially early in training before FB cues develop inhibitory properties. However, conditioned inhibition likely develops faster than is currently appreciated. For instance, we have observed evidence for summation in as little as three trials with an explicitly unpaired fear conditioning procedure (reference[78], Fig. S1)

## Methods

**Rats.** All procedures were performed in accordance with the National Institute of Health Guide for the Care and Use of Experimental Animals and were approved by the Nathan Kline Institute (NKI) and Texas A&M University (TAMU) Animal Care and Use Committees. Most studies were conducted at NKI with a subset conducted at TAMU (Fig. 6 & Extended Fig. 6). Subjects were female and male Sprague-Dawley adult rats weighing 250-300g upon arrival (NKI: Hilltop Laboratory Animals, Scottsdale, PA, USA; TAMU: Envigo, Indianapolis, IN, USA). Rats were pair-housed in transparent plastic high-efficiency particulate absorption (HEPA)-filtered cages with cob bedding on a 12:12h (NKI) or (14:10) light-dark cycle within temperature and humidity-controlled environments. Food and water were available ad libitum throughout the duration of the experiment. Animals were allowed a minimum of five days to acclimate to the vivarium before all experiments. Following surgery, animals were singly housed for a minimum of one week and returned to pair-housing once recovered.

**Shuttlebox apparatus.** All equipment was from Coulbourn Instruments. SigAA training and devaluation tests occurred in two-way shuttle boxes (Model H10-11R-SC) within sound-attenuating chambers (SAC Model: H10-24A). Shuttle chambers were divided into two identical sections (25.4 depth x 50.8 width x 30.5 cm height) by metal dividers which contained a small passage (8.0 cm wide, 8.5 cm high) that allowed rats to move freely between sides. In one experiment, the passage connecting the shuttlebox chambers was blocked (door closed) during counterconditioning and one Tone Fear Test via a metal insert held in place with magnets. A houselight with small 0.5 W bulbs in each compartment dimly illuminated the shuttle-boxes (25 cm high). A PC computer installed with Graphic State software (v. 3 or 4) and connected to the chambers via the Habitest Linc System (Model H01-01) delivered 80dB white noise warning stimuli (WSs), feedback (FB) tones (5kHz, 80dB) and shock unconditioned stimuli (USs) during behavioral sessions. WS and FB were produced by an audio generator (Model A12-33) and presented through speakers (Model H12-01R) located on opposing walls in each chamber (20 cm high). Scrambled hock (Model H13-15) was delivered to a grid floor consisting of 32 stainless steel bars arranged in parallel to the dividers (Model H10-11R-XX-SF). Shuttling was automatically registered by infrared beams and recorded throughout the sessions (Model H20-95-X). Behavior was monitored by security video cameras mounted on the ceiling of each shuttlebox chamber and recorded to DVD.

**Counterconditioning apparatus.** All equipment was from Coulbourn Instruments. Counterconditioning and some Tone Fear Tests took place in rodent conditioning chambers (28.5 depth x 26 width x 28.5 height cm, Model H10-11R-TC) enclosed by sound-attenuating cubicles (Model H10-24A, Coulbourn Instruments) and dimly lit by a houselight with a 0.5 W bulb (25 cm high). A computer, installed with Graphic State software (v. 3) and connected to the chambers via the Habitest Linc System (Model H01-01) delivered tones (5s, 5kHz, 80dB) and shock US stimuli during behavioral sessions. Tones were produced by an audio generator (Model A12-33) and presented through a single speaker (Model H12-01R) located in the front panel of the chamber (20 cm high). Scrambled shocks (Model H13-15) were delivered to a grid floor containing 18 stainless steel bars arranged in parallel (H10-11R-TC-SF). Animal movement was monitored by a security video camera mounted on the ceiling of the cubicle and recorded to DVD.

**2-way SigAA training procedure.** Each session began in illuminated shuttleboxes with a 5-minute stimulus-free acclimation. On the first day of training only, after acclimation, rats received a single Pavlovian (WS-US paired) trial with an inescapable 0.5s footshock US, followed by 30 SigAA trials. The remaining days of training (day 2+) did not include the initial Pavlovian trial. After the acclimation period on subsequent training days (day 2+), rats received 30 SigAA trials with an average intertrial interval

(ITI) of 2 minutes (random assignment without replacement, range 90-150s). Trials consisted of an 80dB white noise WS presentation that lasted a maximum of 15 s. Shuttling during the WS led to immediate WS termination, prevented delivery of the scheduled footshock US (0.5 s; 1.0 mA males, 0.7 mA females), and triggered the delivery of a pure tone FB cue (5s, 5kHz, 80dB), or in the case of FB-controls (see below), a 5s period of darkness (houselights turned off). If no avoidance shuttle was performed during the 15s WS, the inescapable US was delivered for 0.5s and no FB was delivered. Avoidance responses (ARs) were defined as shuttles during the WS. After 30 avoidance trials were completed, animals were given a 60s final period in the shuttle box before removal. Training (one session/day) continued for either 5 days (moderately trained) or 20 days (overtrained) with 5 sessions occurring per week. Rats failing to reach 70% successful avoidance by session 5 (poor avoiders) were removed from experiments.

**Counterconditioning procedure.** During the last five days of avoidance training (days 1-5 for moderately-trained rats, days 16-20 for overtrained rats), animals were placed in the illuminated counterconditioning apparatus and given 30 min. stimulus-free habituation sessions. It is our impression that this habituation phase is critical for obtaining effective counterconditioning of safety signals. 1d after the final SigAA training session, animals were placed in the illuminated counterconditioning boxes for a 3 min. acclimation period. FB-Paired and FB-control rats received 40 trials of the 5s tone (5kHz, 80dB) that co-terminated with a 0.5s footshock US (0.7mA for females, 1.0mA for males; avg. ITI: 180s, random assignment without replacement, range 150- 210 s). For the FB-control group, the tone was novel since they had been SigAA-trained with a visual FB cue. Unpaired rats received 40 tones and shocks in an explicitly unpaired fashion. Trial 1 began with a footshock, with tones and shocks alternating thereafter. Shocks occurred on average every 3 minutes, with placement closer to the onset of the next tone than the termination of the previous tone (to prevent trace conditioning). Shock-to-tone intervals ranged from 28-88s (mean: 58s) and tone-to-shock intervals ranged from 87-147s (mean: 117s). All intervals were selected from lists at random without replacement. Total session times for all animals were approximately 125 min. Two experiments with female rats conducted counterconditioning in the shuttleboxes using only 20 tones & shocks (Paired versus Unpaired, Extended Fig. 5). These sessions were the same as the standard procedure except the total session duration was shorter (~65 min).

**Devaluation/AR Tests.** Note that all Devaluation or AR Tests were conducted 10 min. after i.v. injections with 1 ml/kg injection volumes. For within-subject KORD studies, rats were injected with Veh or SalB (counterbalanced). For other purely behavioral experiments, rats were injected with Veh before both tests. This was done to control for stress before testing in all studies, and to allow team members to practice i.v. injections before the critical KORD studies. After acclimation in the shuttleboxes (SigAA training context) for 5 min., rats received 15 presentations of the 80dB, 15s WS (mean ITI: 120s, random assignment without replacement, range 90-150s). Shuttling was monitored throughout the session but had no effect on WS duration (no WS-termination). There were also no FB presentations (tones or houselight-off) or footshock US presentations. Devaluation/AR tests occurred twice, with 1d off between tests.

**Tone fear tests.** After all devaluation tests were completed, rats from behavioral experiments were placed in a novel open field (90 x 90 x 40 cm) for a 2 minute acclimation phase followed by 3 presentations of a 5 kHz, 80 dB, 60s tone cue (1 min ITI) (Figs. 2d, 2j, 5d, 5j, and Extended Figs. 5f,l,r). Video was recorded to DVD using a digital camera and freezing responses were rated offline. Some rats were also tested in the two-way shuttleboxes where the avoidance acquisition training originally occurred using an identical protocol with the shuttlebox door open or closed (Figs. 2j, 5j, and

Extended Fig. 5f,l,r). Some rats were subsequently tested in the counterconditioning boxes (Figs. 2j, 5j) using an identical protocol. Tone fear tests occurred 2-3d after the final devaluation/AR Test, with only one test/day (2-3d interval between tests).

**USAA procedure.** Rats were placed in the shuttlebox, with offset of the houselight indicating the beginning of the session. Shock (1.0 mA for males, 0.7 mA for females) was delivered every 5s (shock-shock, or S-S interval) unless the subject shuttled to the opposite chamber. Every shuttle earned a FB tone (0.5 s, 70 dB, 2kHz) and a 30s shock-free period (response-shock, or R-S interval). Shuttles during S-S intervals were considered escape responses (ERs) and shuttles during R-S intervals were considered avoidance responses (ARs). If the subject failed to shuttle during S-S or R-S intervals, shocks resumed every 5 s. All subjects received USAA until reaching “good avoider criterion” (2 consecutive days with 20 or more avoidance responses) [27]. Rats failing to reach this criterion by session 10 were deemed poor avoiders and eliminated from the study. Session duration was 25 minutes.

Contingency Degradation (CD) phase: Following USAA training, animals were nonrandomly assigned to Control or CD groups based on comparable levels of criterion performance. Control animals received four more days of USAA training, while experimental animals received four days of USAA-CD. For the latter, all parameters were the same as the USAA protocol except during the shock-free period (regardless of whether it was earned by escape or avoidance), every 5s there was a 1/3 chance that the rat would receive a ‘free’ SS (and additional shock-free period) in a way that was not contingent on behavior.

**Surgery for DREADD experiments.** Rats were anesthetized with a mixture of ketamine (100 mg/kg, i.p.) and xylazine (10 mg/kg, i.p.) and placed in a stereotaxic apparatus (David Kopf Instruments). Supplemental doses of the mixture were given as needed to maintain a deep level of anesthesia. Brain areas were targeted using coordinates from the Paxinos and Watson Brain Atlas. The skull was exposed and small burr holes were drilled above the targeted area. Rats were bilaterally injected with 0.5 uL AAV using a 1.0 uL Hamilton Neuros syringe (Hamilton Co.) at a rate of 0.1 uL/min. Following infusion, the syringe was left in place for an additional 5-8 minutes to allow fluid diffusion away from the needle tip before withdrawing. AAV injections consisted of a 1:10 cocktail of KORD [AAV8-HSyn-DIO-HA-KORD-IRES-mCitrine (Addgene plasmid no. 65417), packaged at UPenn Vector Core, 8.9 × 1013GC ml−1] and Cre [AAV5-CMV-HI-eGFP-Cre.WPRE.SV40 (Addgene plasmid no. 105545) packaged at UPenn Vector Core 2.5 × 1014 GC ml−1], respectively. Injections were made into the anterior dorsomedial striatum (aDMS; y = +0.5 mm anterior to Bregma, x = +/- 2.1 mm lateral to midline, z = -4.5 mm ventral from skull surface), the posterior dorsomedial striatum (pDMS; y = -0.12 mm, x = +/- 2.45 mm, z = -4.9 mm), or the dorsolateral striatum (DLS; y = +0.75 mm, x = +/- 3.5 mm, z = -5.1, -5.6 mm). Immediately following surgery, animals were sutured (Ethicon Vicryl, Fisher Scientific), injected with analgesic (Buprenorphine-ER (s.c., 0.6-1.2 mg/kg), Wedgewood Pharmacy) and allowed to recover on a heat pad. Once awake and alert, rats were returned to the vivarium in a fresh, clean cage to recover for one week singly housed. After one week, rats were returned to pair housing for at least 14 additional days before behavioral training.

**Drug injections.** Salvinorin B (SalB, 2.5 mg/mL, Sigma) was dissolved in Vehicle (7% DMSO and 50% PEG in ddH2O) under low heat. Animals were restrained briefly and intravenously injected (i.v., lateral tail vein) with Vehicle or 2.5 mg/kg SalB (1ml/kg volume) 10 min prior to all devaluation/AR Tests. Treatments (Vehicle versus SalB) and order of treatment (Vehicle first versus SalB first) were counterbalanced for all KORD studies evaluating DMS and DLS. Animals with poor or unsuccessful injections were removed from the study.

**Implant surgery for electrophysiology.** 3-4 weeks after unilateral DMS viral transfection using techniques and coordinates described above, 4 rats were surgically anesthetized with isoflurane, and a topical local anesthetic (1% xylocaine) was administered locally around the wound site. Temperature was maintained at 37°C throughout the surgery using a heating pad. Animals were placed in a stereotaxic device and the skull was exposed. A 2-channel telemetry transmitter (HD-X02, Data Sciences International) was inserted under the skin of the back and stainless-steel electrodes (100μ diameter, Teflon coated, A-M Systems) connected to the telemetry leads were implanted in the DMS bilaterally targeting the region of KORD transfection in the infected hemisphere and the same location in the contralateral uninfected hemisphere. Reference electrodes were implanted in the posterior cortex. Electrodes were attached to the skull with bone-anchor screws and dental cement and the scalp sutured closed. Following recovery from anesthesia, animals were single housed and allowed to recover for at least 1 week before recording.

***In vivo* recording and analysis.** Recordings were performed in the animal's home cage by placing the cage on a telemetry receiver. Bilateral local field potentials were recorded in the DMS (amplification 200x, low-pass filtered at 200Hz) and digitized (1000Hz) for recording and analysis with Spike2 software (CED, Inc). Following at least 15 min of baseline recording, rats were injected with SalB (2.5 mg/kg, 1ml/kg, i.v.) and returned to their home cage (<5 min post-injection) for continued recording of spontaneous activity for at least 30min. LFP data were analyzed with FFT (2.4Hz/bin) from at least 10 min. of continuous data pre-injection and starting at least 15 min. post-injection. The post-injection FFT spectrum was normalized to the baseline recording from that same hemisphere, providing a measure of SalB induced change in the infected DMS and the control, uninfected DMS within animals. Following the recording, animals were perfused for histological confirmation of viral expression and electrode placement. In one animal the electrode in the control hemisphere missed the DMS and entered the lateral ventricle, thus only data from the infected hemisphere of this animal was used in the analysis (Extended Data Fig. 3a-d).

**Histology and immunohistochemistry.** After behavioral testing was completed, DMS and DLS animals were anaesthetized with an overdose mixture of ketamine (100 mg/kg, i.p.) and xylazine (10 mg/kg, i.p.) and perfused transcardially with 4% paraformaldehyde (PFA) in 0.01 M phosphate-buffered saline (PBS). The brains were fixed in 4% PFA solution for 24 hours, then post-fixed in 30% sucrose at 4ºC for 48-72 hours. Brains were then blocked and sectioned on a microtome (SM2000R, Leica) at 40 um. These sections were kept in 0.01 M PBS + 0.05% sodium azide (NaAz) until immunohistochemistry. Brain sections were blocked in 1% bovine serum albumin (BSA; Sigma) in 0.1 M PBS for 30-60 minutes at room temperature (RT) to block non-specific binding. KORD-injected brain sections were incubated overnight at RT in Rabbit Anti-GFP Polyclonal Primary Antibody (Invitrogen, #A-11122) in 1% BSA + 0.2% Triton X-100 in 0.01 M PBS (1:1,000) for verification of viral expression in neurons. Following primary antibody incubation, sections were rinsed with agitation three times for 5 minutes each in 0.01 M PBS at RT. Sections were then incubated in Goat Anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 488 (Invitrogen, #A-11008) in 0.01 M PBS (1:200) at RT for 30-60 minutes. In some experiments primary and secondary antibody steps were combined using Alexa Fluor 488-tagged GFP Polyclonal Antibody, (Invitrogen, #A21311). After secondary incubation, sections were rinsed with agitation three times for 5 minutes each in 0.01 M PBS at RT, mounted on gelatin-coated slides, cover-slipped in Permount (Fisher, #SP-15) or Prolong Gold Antifade Reagent (Invitrogen, #P36930) and allowed to cure overnight at RT. Slides were placed at 4ºC for long-term storage. Sections were imaged using a Leica TCS SP8 confocal microscope (Leica) or Olympus VS120 fluorescent microscope (Olympus). For all experiments, animals were excluded from analysis if virus expression was insufficient, too extensive, or outside regions of interest.

**Assessment of freezing.** Videos recorded during counterconditioning and testing sessions were analyzed for freezing by at least one rater who was blind to treatment condition when possible. Blinding was not possible for counterconditioning sessions where reactions to shock were obvious. Freezing was defined as cessation of all non-respiratory movement and is presented as a percentage of the interval assessed. For counterconditioning sessions, freezing was scored for the first 5 trials during pre-tone (5s) or tone (Paired: 4.5s, Unpaired: 5s) intervals. For Devaluation and AR Test sessions, freezing was scored during 15s WS presentations for the first 3 trials (Block 1). For Tone Fear Tests, freezing was scored during each 60s tone presentation.

**Statistics and reproducibility.** Data are typically presented as mean ± s.e.m, or mean plus individual values. Most data were analyzed using GraphPad Prism (v. 10.0.3) software. Figures were created with GraphPad Prism, Adobe Illustrator, and Biorender.com. Error bars in graphs represent ± s.e.m for between-subject analyses, error bars are not included for within-subject graphs. Unpaired Students t-tests were used to compare means between two groups. Paired Students t-tests were used for within-subject comparisons at a single timepoint (e.g. Block 1 histograms). For comparing multiple groups at a single timepoint, ordinary One-way ANOVAs were conducted, assuming equal SDs per group, followed by Dunnett's multiple comparisons test (single pooled variance). For comparing groups across multiple timepoints, Two-way repeated measures ANOVAs were performed with no assumption of sphericity (Geisser-Greenhouse correction). For experiments with more than two groups, this was followed by Dunnett's multiple comparison test, with individual variances computed for each comparison. For experiments with two groups, this was followed by Šídák's multiple comparison test, with individual variances computed for each comparison. To determine variance among animals across sessions (Fig. 1d, right panel), eta-squared ($\eta^2$) was calculated to determine effect size, where $\eta^2 \leq 0.06$, 0.07–0.14, and $\eta 2 > 0.14$ indicated small, medium, and large effect sizes, respectively. For each animal, a significant large sized session effect ($F = 8.37$, $p < 0.001$, $\eta^2 = 0.17$) was apparent. We therefore calculated coefficients of variation (CV) to assess the extent of variability across sessions was significant using a mixed model ANOVA. Here, CV is inversely related to stereotyped behavior. Effect size calculations were completed using R statistical software. P values <0.05 were considered statistically significant.

## Data availability

All the data that support the findings presented in this study are available from the corresponding author upon reasonable request.

## Acknowledgments

The authors thank Andrew Delamater, Vincent Campese, and Cristina Siller-Perez for helpful discussions on the design of the devaluation procedure and interpretation of data. The project described was supported by Award Number R01MH114931 to C.K.C. from the U.S. National Institutes of Health. R.M.S. received support from a Templeton Foundation grant (TWCF0366). The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institute of Health and National Institute of Mental Health.

## Author contributions

RM conducted surgeries and histology, analyzed and interpreted data, and assisted with writing. EA conducted surgeries, histology, perfusions, and data analyses. SS, LL, and DM participated in all aspects of data collection including surgeries, perfusions, histology, behavioral testing, and scoring of freezing. DW conducted the *in vivo* electrophysiology experiment and analyzed data. MA conducted the contingency degradation experiment at TAMU. JM designed the contingency degradation experiment, collected and analyzed behavioral data, and assisted with writing. CC designed the devaluation experiments, collected and analyzed behavioral data, and wrote the paper.

## References


1 LeDoux, J. E. & Gorman, J. M. A call to action: overcoming anxiety through active coping. *Am J Psychiatry* **158**, 1953-1955. (2001).

2 Collins, K. A., Mendelsohn, A., Cain, C. K. & Schiller, D. Taking action in the face of threat: neural synchronization predicts adaptive coping. *J Neurosci* **34**, 14733-14738, doi:10.1523/JNEUROSCI.2152-14.2014 (2014).

3 Lovibond, P. F., Mitchell, C. J., Minard, E., Brady, A. & Menzies, R. G. Safety behaviours preserve threat beliefs: Protection from extinction of human fear conditioning by an avoidance response. *Behav Res Ther* **47**, 716-720, doi:10.1016/j.brat.2009.04.013 (2009).

4 Gillan, C. M. *et al.* Enhanced avoidance habits in obsessive-compulsive disorder. *Biol Psychiatry* **75**, 631-638, doi:10.1016/j.biopsych.2013.02.002 (2014).

5 Cain, C. K. Avoidance Problems Reconsidered. *Curr Opin Behav Sci* **26**, 9-17, doi:10.1016/j.cobeha.2018.09.002 (2019).

6 LeDoux, J. E., Moscarello, J., Sears, R. & Campese, V. The birth, death and resurrection of avoidance: a reconceptualization of a troubled paradigm. *Molecular psychiatry* **22**, 24-36, doi:10.1038/mp.2016.166 (2017).

7 Rescorla, R. A. & Solomon, R. L. Two process learning theory: Relationships between Pavlovian conditioning and instrumental learning. *Psychological Review* **74**, 151-182 (1967).

8 Dinsmoor, J. A. Stimuli inevitably generated by behavior that avoids electric shock are inherently reinforcing. *Journal of the Experimental Analysis of Behavior* **75**, 311-333 (2001).

9 Balleine, B. W. & Dickinson, A. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* **37**, 407-419 (1998).

10 Gillan, C. M., Urcelay, G. & Robbins, T. W. in *The Wiley Handbook on the Cognitive Neuroscience of Learning* (eds R.A. Murphy & R.C. Honey) Ch. 17, 442-467 (John Wiley & Sons, LTD, 2016).

11 Rescorla, R. A. Establishment of a positive reinforcer through contrast with shock. *J Comp Physiol Psychol* **67**, 260-263, doi:10.1037/h0026789 (1969).

12 Soltysik, S. Inhibitory Feedback in Avoidance Conditioning. *Bol Inst Estud Med Biol Univ Nac Auton Mex* **21**, 433-449 (1963).

13 Bolles, R. C. in *The Psychology of Learning and Motivation* Vol. 6 (eds G.H. Bower & J.T. Spence) 97-145 (Academic Press, 1972).

14 Weisman, R. G. & Litner, J. S. Positive conditioned reinforcement of Sidman avoidance behavior in rats. *J Comp Physiol Psychol* **68**, 597-603 (1969).

15 Candido, A., Gonzalez, F. & de Brugada, I. Safety signals from avoidance learning but not from yoked classical conditioning training pass both summation and retardation tests for inhibition. *Behavioural processes* **66**, 153-160, doi:10.1016/j.beproc.2004.01.011 (2004).

16 Morris, R. G. M. Preconditioning of reinforcing properties to an exteroceptive feedback stimulus. *Learning and Motivation* **6**, 289-298 (1975).

17 Fernando, A. B., Urcelay, G. P., Mar, A. C., Dickinson, A. & Robbins, T. W. Safety signals as instrumental reinforcers during free-operant avoidance. *Learn Mem* **21**, 488-497, doi:10.1101/lm.034603.114 (2014).

18 Yin, H. H., Ostlund, S. B., Knowlton, B. J. & Balleine, B. W. The role of the dorsomedial striatum in instrumental conditioning. *Eur J Neurosci* **22**, 513-523, doi:10.1111/j.1460-9568.2005.04218.x (2005).

19 Adams, C. D. & Dickinson, A. Instrumental responding following reinforcer devaluation. *Q J Exp Psychol (Colchester)*, 109-121 (1981).

20 Yin, H. H., Knowlton, B. J. & Balleine, B. W. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur J Neurosci* **19**, 181-189 (2004).

21 Adams, C. D. Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Q J Exp Psychol* **34**, 77-98 (1982).
22 Garr, E. & Delamater, A. R. Exploring the relationship between actions, habits, and automaticity in an action sequence task. *Learn Mem* **26**, 128-132, doi:10.1101/lm.048645.118 (2019).
23 Balleine, B. W. & Dezfouli, A. Hierarchical Action Control: Adaptive Collaboration Between Actions and Habits. *Frontiers in psychology* **10**, 2735, doi:10.3389/fpsyg.2019.02735 (2019).
24 Rescorla, R. A. Pavlovian conditioned inhibition. *Psychological bulletin* **72**, 77-94 (1969).
25 Declercq, M. & De Houwer, J. On the role of US expectancies in avoidance behavior. *Psychonomic bulletin & review* **15**, 99-102 (2008).
26 Ramirez, F., Moscarello, J. M., LeDoux, J. E. & Sears, R. M. Active avoidance requires a serial basal amygdala to nucleus accumbens shell circuit. *J Neurosci* **35**, 3470-3477, doi:10.1523/JNEUROSCI.1331-14.2015 (2015).
27 Lazaro-Munoz, G., LeDoux, J. E. & Cain, C. K. Sidman instrumental avoidance initially depends on lateral and basal amygdala and is constrained by central amygdala-mediated Pavlovian processes. *Biol Psychiatry* **67**, 1120-1127 (2010).
28 Moscarello, J. M. & Penzo, M. A. The central nucleus of the amygdala and the construction of defensive modes across the threat-imminence continuum. *Nat Neurosci* **25**, 999-1008, doi:10.1038/s41593-022-01130-5 (2022).
29 Vardy, E. *et al.* A New DREADD Facilitates the Multiplexed Chemogenetic Interrogation of Behavior. *Neuron* **86**, 936-946, doi:10.1016/j.neuron.2015.03.065 (2015).
30 Thorn, C. A. & Graybiel, A. M. Differential entrainment and learning-related dynamics of spike and local field potential activity in the sensorimotor and associative striatum. *J Neurosci* **34**, 2845-2859, doi:10.1523/JNEUROSCI.1782-13.2014 (2014).
31 DeCoteau, W. E. *et al.* Oscillations of local field potentials in the rat dorsal striatum during spontaneous and instructed behaviors. *J Neurophysiol* **97**, 3800-3805, doi:10.1152/jn.00108.2007 (2007).
32 Tallot, L., Graupner, M., Diaz-Mataix, L. & Doyere, V. Beyond Freezing: Temporal Expectancy of an Aversive Event Engages the Amygdalo-Prefronto-Dorsostriatal Network. *Cereb Cortex* **30**, 5257-5269, doi:10.1093/cercor/bhaa100 (2020).
33 Balleine, B. W., Liljeholm, M. & Ostlund, S. B. The integrative function of the basal ganglia in instrumental conditioning. *Behav Brain Res* **199**, 43-52, doi:10.1016/j.bbr.2008.10.034 (2009).
34 Lipton, D. M., Gonzales, B. J. & Citri, A. Dorsal Striatal Circuits for Habits, Compulsions and Addictions. *Frontiers in systems neuroscience* **13**, 28, doi:10.3389/fnsys.2019.00028 (2019).
35 Wendler, E. *et al.* The roles of the nucleus accumbens core, dorsomedial striatum, and dorsolateral striatum in learning: performance and extinction of Pavlovian fear-conditioned responses and instrumental avoidance responses. *Neurobiol Learn Mem* **109**, 27-36, doi:10.1016/j.nlm.2013.11.009 (2014).
36 Peak, J., Hart, G. & Balleine, B. W. From learning to action: The integration of dorsal striatal input and output pathways in instrumental conditioning. *Eur J Neurosci*, doi:10.1111/ejn.13964 (2018).
37 Corbit, L. H., Ostlund, S. B. & Balleine, B. W. Sensitivity to instrumental contingency degradation is mediated by the entorhinal cortex and its efferents via the dorsal hippocampus. *J Neurosci* **22**, 10976-10984 (2002).
38 Bolles, R. C. & Popp, R. J., Jr. Parameters affecting the acquisition of Sidman avoidance. *J Exp Anal Behav* **7**, 315-321 (1964).
39 Rescorla, R. A. Pavlovian conditioned fear in Sidman avoidance learning. *J Comp Physiol Psychol* **65**, 55-60 (1968).
40 Fanselow, M. S. in *Learning, Motivation, and Cognition: The Functional Behaviorism of Robert C. Bolles* (eds M. E. Bouton & M. S. Fanselow) Ch. 16, 321-341 (American Psychological Association, 1997).

41 Krypotos, A. M., Effting, M., Kindt, M. & Beckers, T. Avoidance learning: a review of theoretical models and recent developments. *Front Behav Neurosci* **9**, 189, doi:10.3389/fnbeh.2015.00189 (2015).
42 Dickinson, A. in *Associative learning and cognition: homage to professor N. J. Mackintosh, "in memoriam" (1935-2015)* (eds J. B. Trobalon & V. D. Chamizo) 177-196 (Edicions de la Universitat de Barcelona, 2016).
43 Yin, H. H., Knowlton, B. J. & Balleine, B. W. Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. *Eur J Neurosci* **22**, 505-512, doi:10.1111/j.1460-9568.2005.04219.x (2005).
44 Choi, J. S., Cain, C. K. & LeDoux, J. E. The role of amygdala nuclei in the expression of auditory signaled two-way active avoidance in rats. *Learn Mem* **17**, 139-147 (2010).
45 Poremba, A. & Gabriel, M. Amygdala neurons mediate acquisition but not maintenance of instrumental avoidance behavior in rabbits. *J Neurosci* **19**, 9635-9641 (1999).
46 Bravo-Rivera, C., Roman-Ortiz, C., Brignoni-Perez, E., Sotres-Bayon, F. & Quirk, G. J. Neural structures mediating expression and extinction of platform-mediated avoidance. *J Neurosci* **34**, 9736-9742, doi:10.1523/JNEUROSCI.0191-14.2014 (2014).
47 Wenzel, J. M. *et al.* Phasic Dopamine Signals in the Nucleus Accumbens that Cause Active Avoidance Require Endocannabinoid Mobilization in the Midbrain. *Current biology : CB* **28**, 1392-1404 e1395, doi:10.1016/j.cub.2018.03.037 (2018).
48 Stelly, C. E. *et al.* Pattern of dopamine signaling during aversive events predicts active avoidance learning. *Proc Natl Acad Sci U S A* **116**, 13641-13650, doi:10.1073/pnas.1904249116 (2019).
49 Oleson, E. B., Gentry, R. N., Chioma, V. C. & Cheer, J. F. Subsecond dopamine release in the nucleus accumbens predicts conditioned punishment and its successful avoidance. *J Neurosci* **32**, 14804-14808, doi:10.1523/JNEUROSCI.3087-12.2012 (2012).
50 Fernando, A. B., Urcelay, G. P., Mar, A. C., Dickinson, T. A. & Robbins, T. W. The role of the nucleus accumbens shell in the mediation of the reinforcing properties of a safety signal in free-operant avoidance: dopamine-dependent inhibitory effects of d-amphetamine. *Neuropsychopharmacology* **39**, 1420-1430, doi:10.1038/npp.2013.337 (2014).
51 Darvas, M., Fadok, J. P. & Palmiter, R. D. Requirement of dopamine signaling in the amygdala and striatum for learning and maintenance of a conditioned avoidance response. *Learn Mem* **18**, 136-143, doi:10.1101/lm.2041211 (2011).
52 Soltysik, S. The avoidance reflex reinforced with food. *Acta Biologiae Experimentalis* **29** (1969).
53 Bower, G., Starr, R. & Lazarovitz, L. Amount of Response-Produced Change in the Cs and Avoidance Learning. *J Comp Physiol Psychol* **59**, 13-17, doi:10.1037/h0021623 (1965).
54 Dinsmoor, J. A. & Sears, G. W. Control of avoidance by a response-produced stimulus. *Learn Motiv* **4**, 284-293 (1975).
55 Braud, W. G. Diminution of suppression by stimuli associated with the offset of fear-arousing cues. *J Comp Physiol Psychol* **65**, 356-358, doi:10.1037/h0025554 (1968).
56 Hendersen, R. W. & Graham, J. Avoidance of heat by rats: effects of thermal context on rapidity of extinction. *Learning and Motivation* **10**, 351-363 (1979).
57 de Wit, S. *et al.* Shifting the balance between goals and habits: Five failures in experimental habit induction. *J Exp Psychol Gen* **147**, 1043-1065, doi:10.1037/xge0000402 (2018).
58 Gordon, A. L., Patterson, T. K. & Knowlton, B. J. Early-life stress is associated with a preponderance of habitual responding in a novel instrumental avoidance learning paradigm. *Neurobiol Learn Mem* **175**, 107316, doi:10.1016/j.nlm.2020.107316 (2020).
59 Fernando, A., Urcelay, G., Mar, A., Dickinson, A. & Robbins, T. Free-operant avoidance behavior by rats after reinforcer revaluation using opioid agonists and D-amphetamine. *J Neurosci* **34**, 6286-6293, doi:10.1523/JNEUROSCI.4146-13.2014 (2014).

60 Dickinson, A. & Balleine, B. W. Motivational control of goal-directed action. *Animal Learning and Behavior* **22** (1994).
61 Lysle, D. T. & Fowler, H. Inhibition as a "slave" process: deactivation of conditioned inhibition through extinction of conditioned excitation. *J Exp Psychol Anim Behav Process* **11**, 71-94, doi:10.1037//0097-7403.11.1.71 (1985).
62 Kamin, L. J. The effects of termination of the CS and avoidance of the US on avoidance learning. *J Comp Physiol Psychol* **49**, 420-424 (1956).
63 Bolles, R. C., Stokes, L. W. & Younger, M. S. Does CS termination reinforce avoidance behavior? *Journal of Comparative and Physiological Psychology* **62**, 201-207 (1966).
64 Solomon, R. L. & Wynne, L. C. Traumatic avoidance learning: The principles of anxiety conservation and partial irreversibility. *Psychol. Rev.* **61**, 353 (1954).
65 Masterson, F. A. & Crawford, M. The defense motivation system: A theory of avoidance behavior. *The Behavioral and Brain Sciences* **5**, 661-696 (1982).
66 Declercq, M. & De Houwer, J. Evidence for a hierarchical structure underlying avoidance behavior. *J Exp Psychol Anim Behav Process* **35**, 123-128, doi:10.1037/a0012927 (2009).
67 Moscarello, J. M. & Maren, S. Flexibility in the face of fear: Hippocampal-prefrontal regulation of fear and avoidance. *Curr Opin Behav Sci* **19**, 44-49, doi:10.1016/j.cobeha.2017.09.010 (2018).
68 Mobbs, D., Hagan, C. C., Dalgleish, T., Silston, B. & Prevost, C. The ecology of human fear: survival optimization and the nervous system. *Front Neurosci* **9**, 55, doi:10.3389/fnins.2015.00055 (2015).
69 Perusini, J. N. & Fanselow, M. S. Neurobehavioral perspectives on the distinction between fear and anxiety. *Learn Mem* **22**, 417-425 (2015).
70 Baum, M. Extinction of avoidance responding through response prevention (flooding). *Psychological bulletin* **74**, 276-284 (1970).
71 Peck, C. A. & Bouton, M. E. Context and performance in aversive-to-appetitive and appetitive-to-aversive transfer. *Learning and Motivation* **21**, 1-31 (1990).
72 Binette, A. N., Totty, M. S. & Maren, S. Sex differences in the immediate extinction deficit and renewal of extinguished fear in rats. *PLoS One* **17**, e0264797, doi:10.1371/journal.pone.0264797 (2022).
73 Anderson, L. C. & Petrovich, G. D. Renewal of conditioned responding to food cues in rats: Sex differences and relevance of estradiol. *Physiol Behav* **151**, 338-344, doi:10.1016/j.physbeh.2015.07.035 (2015).
74 Wendt, J., Low, A., Weymar, M., Lotze, M. & Hamm, A. O. Active avoidance and attentive freezing in the face of approaching threat. *NeuroImage* **158**, 196-204, doi:10.1016/j.neuroimage.2017.06.054 (2017).
75 Hashemi, M. M. *et al.* Neural Dynamics of Shooting Decisions and the Switch from Freeze to Fight. *Scientific reports* **9**, 4240, doi:10.1038/s41598-019-40917-8 (2019).
76 Vervliet, B., Lange, I. & Milad, M. R. Temporal dynamics of relief in avoidance conditioning and fear extinction: Experimental validation and clinical relevance. *Behav Res Ther* **96**, 66-78, doi:10.1016/j.brat.2017.04.011 (2017).
77 Lloyd, K. & Dayan, P. Safety out of control: dopamine and defence. *Behav Brain Funct* **12**, 15, doi:10.1186/s12993-016-0099-7 (2016).
78 Ostroff, L. E., Cain, C. K., Bedont, J., Monfils, M. H. & LeDoux, J. E. Fear and safety learning differentially affect synapse size and dendritic translation in the lateral amygdala. *Proc Natl Acad Sci U S A* **107**, 9418-9423 (2010).

Fig. 1

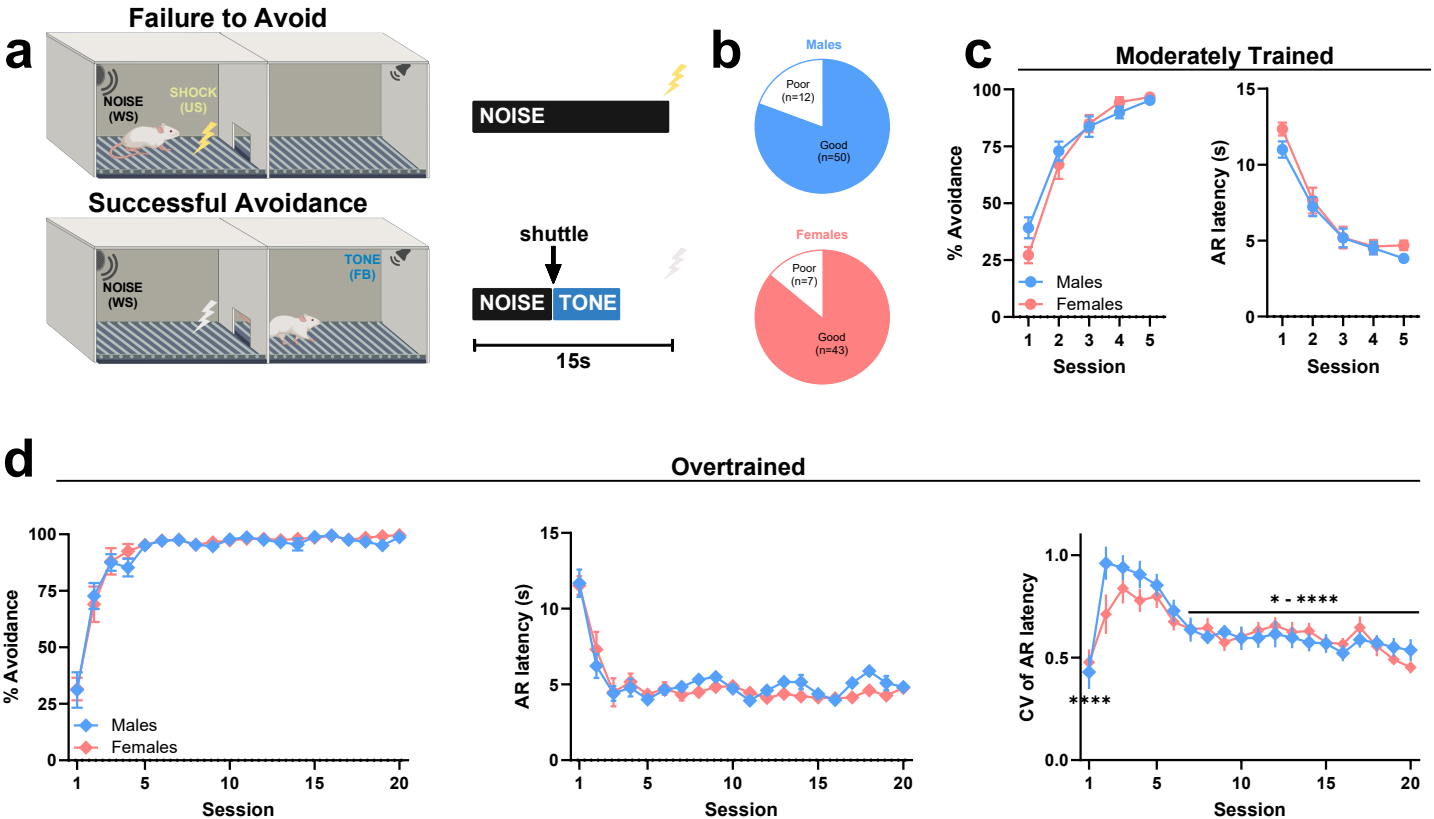

a
Failure to Avoid
NOISE (WS)
SHOCK (US)
Successful Avoidance
NOISE (WS)
TONE (FB)
NOISE
shuttle
NOISE
TONE
15s
b
Males
Poor (n=12)
Good (n=50)
Females
Poor (n=7)
Good (n=43)
c
Moderately Trained
% Avoidance
Males
Females
Session
AR latency (s)
Session
d
Overtrained
% Avoidance
Males
Females
Session
AR latency (s)
Session
CV of AR latency
* - ****
****
Session

**Fig. 1 | Acquisition of SigAA behavior in male and female rats. a**, Schematic of the shuttlebox SigAA paradigm. **b**, Proportion of good versus poor avoiders in males and females from multiple behavioral experiments. **c**, Avoidance behavior during moderate training in male (n=35) and female (n=26) rats. Two-way ANOVA. % Avoidance: $F_{4,236}$=2.28. AR latency: $F_{4,236}$= 0.73. **d**, Avoidance behavior during overtraining in male (n=15) and female (n=17) rats. Two-way ANOVA. % Avoidance: $F_{19,570}$=0.34. AR latency: $F_{19,570}$=0.995. For each animal, a significant large-sized session effect ($F = 8.37$, $p < 0.001$, $\eta^2 = 0.17$) was apparent for AR latency. We therefore calculated coefficients of variation (CV) to assess the extent of variability across sessions using Two-way ANOVA followed by Tukey's multiple comparisons test. Here, CV is inversely related to stereotyped behavior. CV of AR latency: $F_{19,570}$=1.24 (Session x Sex), $F_{19,570}$=11.89 (Session); Session comparisons, Session 5 versus Session 1 ****p<0.0001, Session 5 versus Sessions 7-20 *p=0.02-****p<0.0001. All line graphs are shown as mean ± s.e.m.

Fig. 2

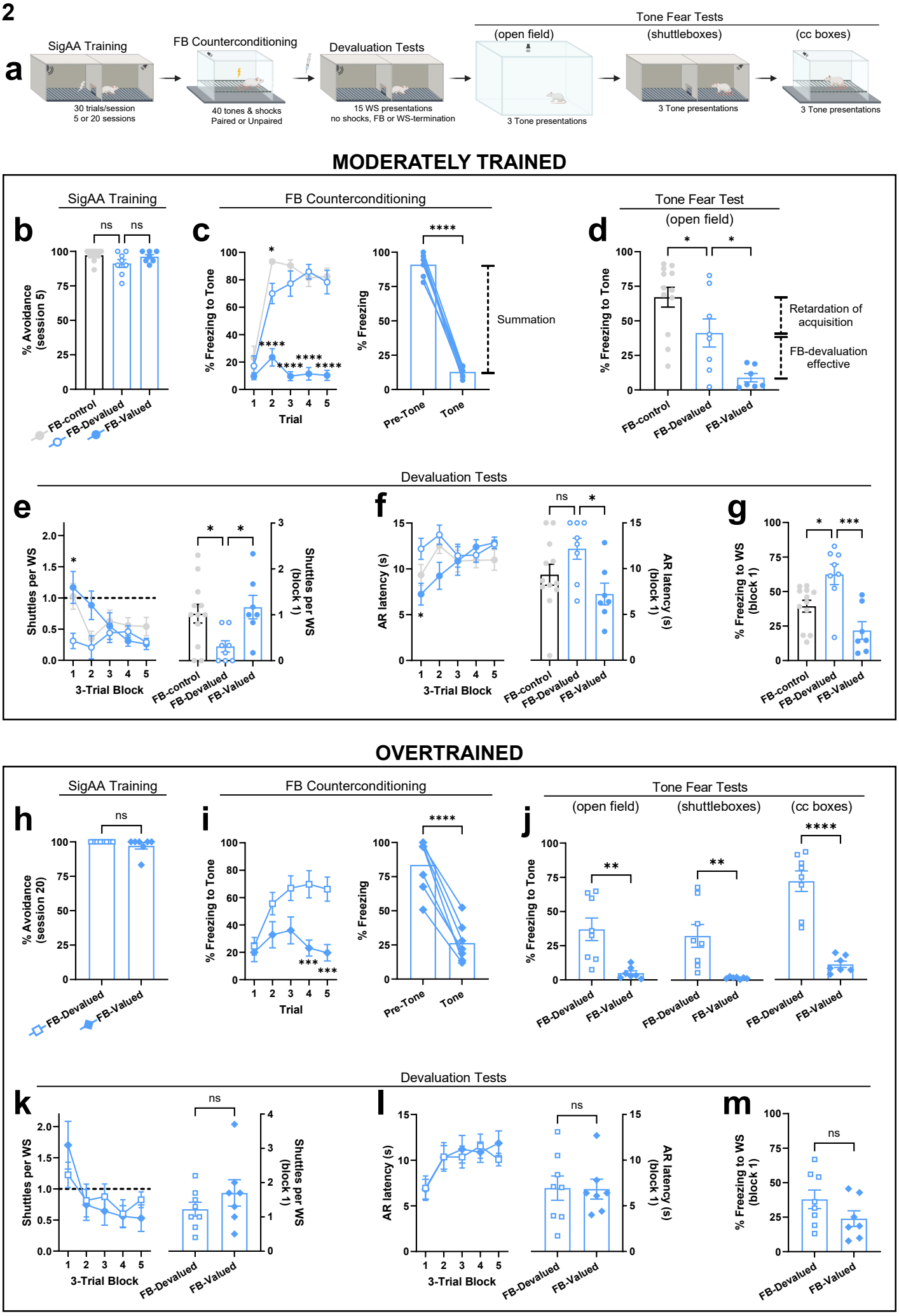

a
SigAA Training
30 trials/session
5 or 20 sessions
FB Counterconditioning
40 tones & shocks
Paired or Unpaired
Devaluation Tests
15 WS presentations
no shocks, FB or WS-termination
Tone Fear Tests
(open field)
3 Tone presentations
(shuttleboxes)
3 Tone presentations
(cc boxes)
3 Tone presentations
MODERATELY TRAINED
SigAA Training
b
% Avoidance (session 5)
ns
ns
FB-control
FB-Devalued
FB-Valued
FB Counterconditioning
c
% Freezing to Tone
Trial
% Freezing
Pre-Tone
Tone
Summation
Tone Fear Test
(open field)
d
% Freezing to Tone
Retardation of acquisition
FB-devaluation effective
Devaluation Tests
e
Shuttles per WS
3-Trial Block
Shuttles per WS (block 1)
f
AR latency (s)
AR latency (s) (block 1)
g
% Freezing to WS (block 1)
OVERTRAINED
SigAA Training
h
% Avoidance (session 20)
FB Counterconditioning
i
% Freezing to Tone
Trial
% Freezing
Pre-Tone
Tone
Tone Fear Tests
(open field)
(shuttleboxes)
(cc boxes)
j
% Freezing to Tone
Devaluation Tests
k
Shuttles per WS
3-Trial Block
Shuttles per WS (block 1)
l
AR latency (s)
AR latency (s) (block 1)
m
% Freezing to WS (block 1)

**Fig. 2 | Devaluation of response-produced safety signals after moderate training versus overtraining in male rats. a**, Schematic of the devaluation paradigm. Panels **b-g**: moderately trained rats (FB-Control: n=12, FB-Devalued: n=8, FB-Valued: n=7). Panels **h-m**: overtrained rats (FB-Devalued: n=8, FB-Valued: 7). Note that moderately trained rats received only one Tone Fear Test in the open field. Additional Tone Fear Tests in the counterconditioning boxes and shuttleboxes were later added for overtrained rats to address differences between males and females. **b**, SigAA performance during session 5 of training. One-way ANOVA. % Avoidance: $F_{2,24}$=2.938. **c**, Left, Tone freezing during the first 5 trials of counterconditioning. Two-way ANOVA followed by Dunnett's test: $F_{8,120}$=5.95. Group comparisons, FB-Devalued versus FB-Control, Trial 2 *p= 0.014; FB-Devalued versus FB-Valued, Trials 2-5 ****p<0.0001. Right, Pre-Tone versus Tone freezing during first 5 trials of counterconditioning for FB-Valued rats. Paired t-test: $t_6$=21.68 ****p<0.0001. **d**, Tone freezing in the novel open field. One-way ANOVA followed by Dunnett's test: $F_{2,24}$=13.95. Group comparisons, FB-control versus FB-Devalued *p=0.04, FB-Devalued versus FB-Valued *p=0.02. **e**, Left, shuttles per WS for both Devaluation Tests. Two-way ANOVA followed by Dunnett's test: $F_{8,96}$=4.41. Group comparisons, FB-Devalued versus FB-control: Trial 1 *p=0.02; FB-Devalued versus FB-Valued: Trial 1 *p=0.03. Right, shuttles per WS for Block 1 of Devaluation Tests. One-way ANOVA followed by Dunnett's test: $F_{2,24}$=4.445. Group comparisons, FB-Devalued versus FB-Control *p=0.03, FB-Devalued versus FB-Valued *p=0.02. **f**, Left, latency to first AR for both Devaluation Tests. Two-way ANOVA followed by test: $F_{8,96}$=3.05. Group comparisons, FB-Devalued versus FB-Valued, Block 1 *p=0.02. Right, latency to first AR during Block 1 of Devaluation Tests. One-way ANOVA followed by Dunnett's test: $F_{2,24}$=3.67. Group comparisons, FB-Devalued versus FB-Control p=0.16, FB-Devalued versus FB-Valued *p=0.02. **g**, WS freezing for Block 1 of Devaluation Tests. One-way ANOVA followed by Dunnett's test: $F_{2,24}$=10.42. Group comparisons, FB-Devalued versus FB-Control *p=0.01, FB-Devalued versus FB-Valued ***p=0.0003. **h**, SigAA performance during session 20 of overtraining for rats assigned to FB-Devalued versus FB-Valued groups. Unpaired t-test: $t_{13}$=1.31 p=0.21. **i**, Left, Tone freezing during the first 5 trials of counterconditioning. Two-way ANOVA followed by Šídák's test: $F_{4,52}$=4.59. Group comparisons, FB-Valued versus FB-Devalued, Trial 4 ***p=0.0009, Trial 5 ***p=0.001. Right, Pre-Tone versus Tone freezing during first 5 trials of counterconditioning for FB-Valued rats. Paired t-test: $t_6$=9.71 ***p<0.0001. **j**, Tone freezing in the novel open field (left), shuttleboxes (middle), and counterconditioning boxes (right). Unpaired t-tests: $t_{13}$=3.58 **p=0.003 (left); $t_{13}$=3.44 **p=0.004 (middle); $t_{13}$=7.27 ****p<0.0001 (right). **k**, Left, shuttles per WS for both Devaluation Tests. Two-way ANOVA: $F_{4,52}$=2.07. Right, shuttles per WS for Block 1 of Devaluation Tests. Unpaired t-test: $t_{13}$=1.14 p=0.28. **l**, Left, latency to first AR for both Devaluation Tests. Two-way ANOVA: $F_{4,52}$=0.81. Right, latency to first AR during Block 1 of Devaluation Tests. Unpaired t-test: $t_{13}$=0.06 p=0.95. **m**, WS freezing for Block 1 of Devaluation Tests. Unpaired t-test: $t_{13}$=1.56 p=0.14. Data in **c** (right) and **i** (right) are shown as mean + individuals. All other data in the figure are shown as mean ± s.e.m.

# Fig. 3

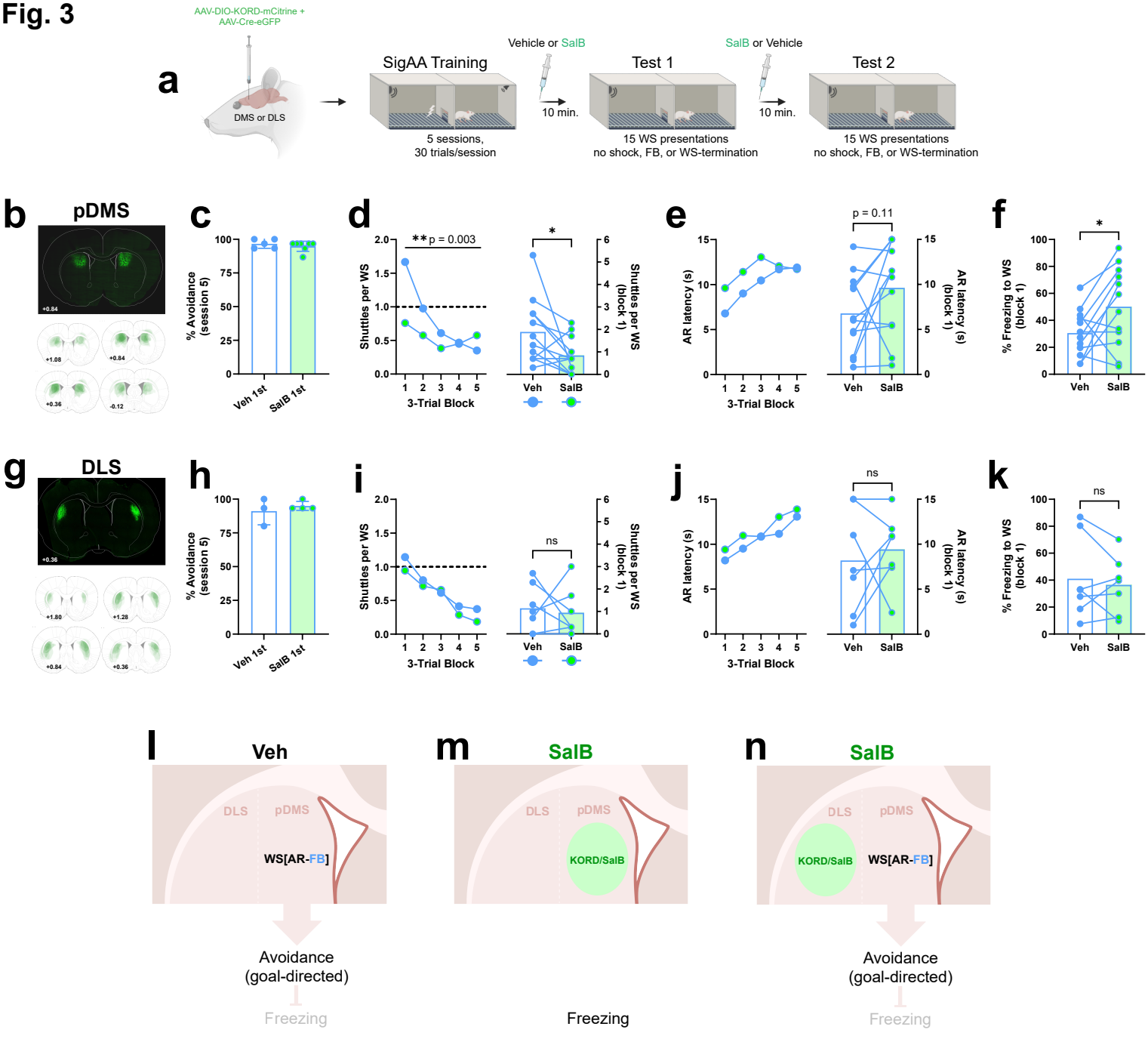

a
AAV-DIO-KORD-mCitrine +
AAV-Cre-eGFP
DMS or DLS
SigAA Training
5 sessions,
30 trials/session
Vehicle or SalB
10 min.
Test 1
15 WS presentations
no shock, FB, or WS-termination
SalB or Vehicle
10 min.
Test 2
15 WS presentations
no shock, FB, or WS-termination
b
pDMS
c
% Avoidance (session 5)
Veh 1st
SalB 1st
d
**p = 0.003
Shuttles per WS
3-Trial Block
*
Shuttles per WS (block 1)
Veh
SalB
e
AR latency (s)
3-Trial Block
p = 0.11
AR latency (s) (block 1)
f
% Freezing to WS (block 1)
*
g
DLS
h
% Avoidance (session 5)
Veh 1st
SalB 1st
i
Shuttles per WS
3-Trial Block
ns
Shuttles per WS (block 1)
j
AR latency (s)
3-Trial Block
ns
AR latency (s) (block 1)
k
% Freezing to WS (block 1)
ns
l
Veh
DLS
pDMS
WS[AR-FB]
Avoidance
(goal-directed)
Freezing
m
SalB
DLS
pDMS
KORD/SalB
Freezing
n
SalB
DLS
pDMS
KORD/SalB
WS[AR-FB]
Avoidance
(goal-directed)
Freezing

**Fig. 3 | Inhibition of pDMS impairs goal-directed avoidance in male rats. a**, Schematic of the experimental paradigm. Panels **b-f**: rats expressing KORD in pDMS (n=12). Panels **g-k**: rats expressing KORD in DLS (n=7). **b**, Top, example of bilateral KORD expression in pDMS. Bottom, Semitransparent overlays illustrating KORD expression in pDMS for individuals across the A-P axis. A-P coordinates are in millimeters relative to Bregma. **c**, SigAA performance during session 5 of training for rats assigned to receive Vehicle versus SalB in Test 1. Unpaired t-test: $t_{10}$=0.89 p=0.40. **d**, Left, shuttles per WS for Vehicle versus SalB Tests. Two-way ANOVA: $F_{4,88}$=4.36 **p=0.003. Right, shuttles per WS for Block 1 of Tests. Paired t-test: $t_{11}$=2.339 *p=0.039. **e**, Left, latency to first AR for Vehicle versus SalB Tests. Two-way ANOVA: $F_{4,88}$=1.85. Right, latency to first AR during Block 1 of Tests. Paired t-test: $t_{11}$=1.713 p=0.115. **f**, WS freezing for Block 1 of Vehicle versus SalB Tests. Paired t-test: $t_{11}$=2.36 p=0.037. **g**, Top, example of bilateral KORD expression in DLS. Bottom, Semitransparent overlays illustrating KORD expression in DLS for individuals across the A-P axis. A-P coordinates are in millimeters relative to Bregma. **h**, SigAA performance during session 5 of training for rats assigned to receive Vehicle versus SalB in Test 1. Unpaired t-test: $t_5$=0.73 p=0.50. **i**, Left, shuttles per WS for Vehicle versus SalB Tests. Two-way ANOVA: $F_{4,48}$=0.11. Right, shuttles per WS for Block 1 of Tests. Paired t-test: $t_6$=0.35 p=0.74. **j**, Left, latency to first AR for Vehicle versus SalB Tests. Two-way ANOVA: $F_{4,48}$=0.22. Right, latency to first AR during Block 1 of Tests. Paired t-test: $t_6$=0.53 p=0.61. **k**, WS freezing for Block 1 of Vehicle versus SalB Tests. Paired t-test: $t_6$=0.66 p=0.53. **l-n**, Diagrammatic summary of results. All data in the figure are shown as means. Error bars (s.e.m.) are shown for between-subject analyses.

Fig. 4

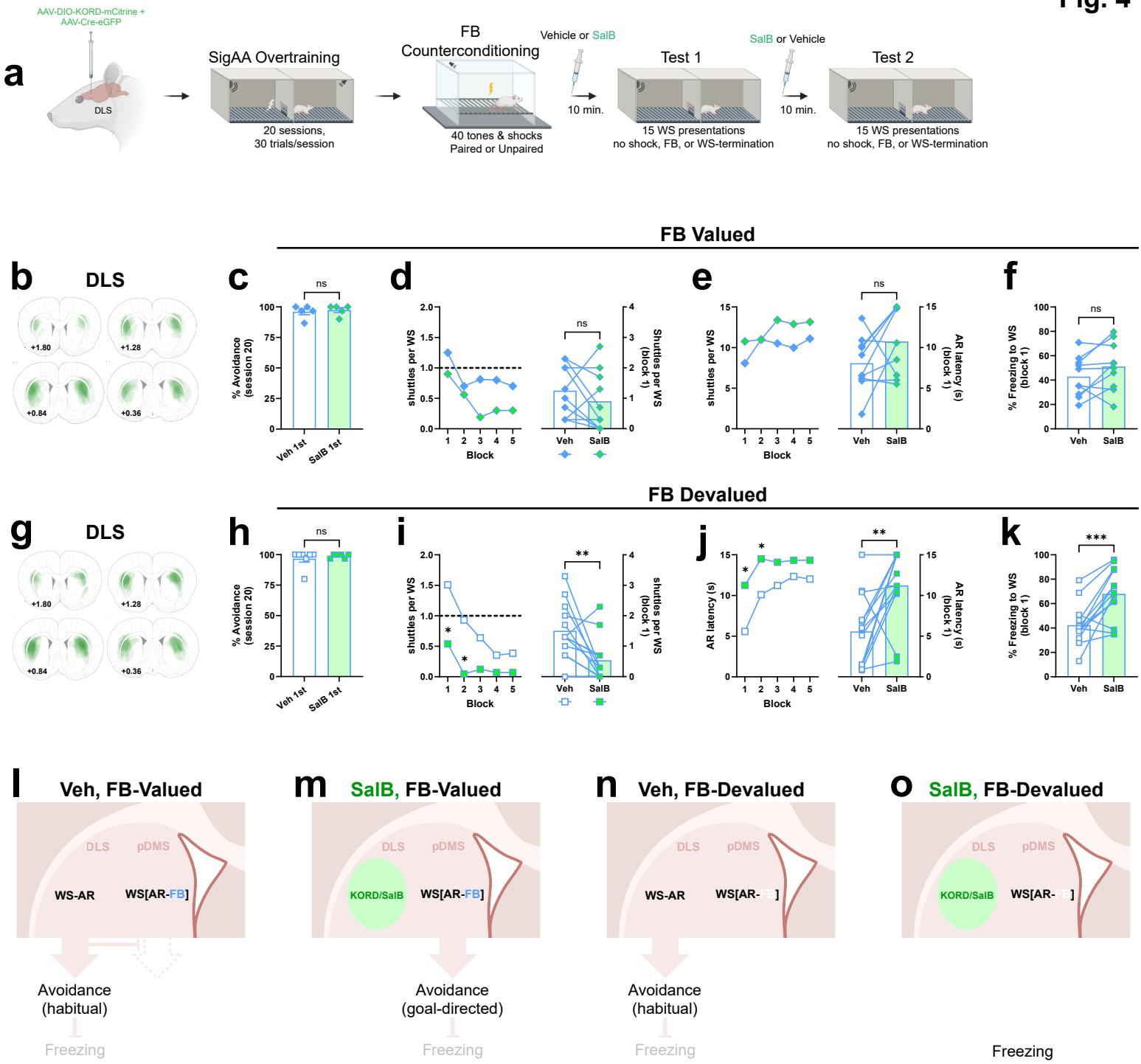

a
AAV-DIO-KORD-mCitrine +
AAV-Cre-eGFP
DLS
SigAA Overtraining
20 sessions,
30 trials/session
FB
Counterconditioning
40 tones & shocks
Paired or Unpaired
Vehicle or SalB
10 min.
Test 1
15 WS presentations
no shock, FB, or WS-termination
SalB or Vehicle
10 min.
Test 2
15 WS presentations
no shock, FB, or WS-termination
FB Valued
b
DLS
+1.80
+1.28
+0.84
+0.36
c
ns
% Avoidance (session 20)
Veh 1st
SalB 1st
d
shuttles per WS
Block
ns
Shuttles per WS (block 1)
Veh
SalB
e
shuttles per WS
Block
ns
AR latency (s) (block 1)
Veh
SalB
f
ns
% Freezing to WS (block 1)
Veh
SalB
FB Devalued
g
DLS
+1.80
+1.28
+0.84
+0.36
h
ns
% Avoidance (session 20)
Veh 1st
SalB 1st
i
shuttles per WS
Block
**
shuttles per WS (block 1)
Veh
SalB
j
AR latency (s)
Block
**
AR latency (s) (block 1)
Veh
SalB
k
***
% Freezing to WS (block 1)
Veh
SalB
l
Veh, FB-Valued
DLS
pDMS
WS-AR
WS[AR-FB]
Avoidance
(habitual)
Freezing
m
SalB, FB-Valued
DLS
pDMS
KORD/SalB
WS[AR-FB]
Avoidance
(goal-directed)
Freezing
n
Veh, FB-Devalued
DLS
pDMS
WS-AR
WS[AR-FB]
Avoidance
(habitual)
Freezing
o
SalB, FB-Devalued
DLS
pDMS
KORD/SalB
WS[AR-FB]
Freezing

**Fig. 4 | Inhibition of DLS impairs habitual avoidance in male rats. a**, Schematic of the experimental paradigm. Note that all rats were overtrained and expressed KORD bilaterally in DLS. All rats received 40 tones and shocks during the counterconditioning phase, either unpaired (FB-Valued, n=10) or paired (FB-Devalued, n=13). **b**, Semitransparent overlays illustrating KORD expression in DLS for individuals assigned to FB-Valued condition. A-P coordinates are in millimeters relative to Bregma. **c**, SigAA performance during session 20 of training for FB-Valued rats assigned to receive Vehicle versus SalB in Test 1. Unpaired t-test: $t_8$=0.42 p=0.68. **d**, Left, shuttles per WS for Vehicle versus SalB Tests in FB-Valued condition. Two-way ANOVA: $F_{4,72}$=0.57. Right, shuttles per WS for Block 1 of Tests in FB-Valued condition. Paired t-test: $t_9$=1.10 p=0.30. **e**, Left, latency to first AR for Vehicle versus SalB Tests for FB-Valued condition. Two-way ANOVA: $F_{4,72}$=0.84. Right, latency to first AR during Block 1 of Tests for FB-Valued condition. Paired t-test: $t_9$=1.76 p=0.11. **f**, WS freezing for Block 1 of Vehicle versus SalB Tests in FB-Valued condition. Paired t-test: $t_9$=1.83 p=0.10. **g**, Semitransparent overlays illustrating KORD expression in DLS for individuals assigned to FB-Devalued condition. A-P coordinates are in millimeters relative to Bregma. **h**, SigAA performance during session 20 of training for FB-Devalued rats assigned to receive Vehicle versus SalB in Test 1. Unpaired t-test: $t_{11}$=0.71 p=0.49. **i**, Left, shuttles per WS for Vehicle versus SalB Tests in FB-Devalued condition. Two-way ANOVA followed by Šídák's test: $F_{4,96}$=2.216 (Block x Drug); $F_{1,24}$=21.23 ***p=0.0001 (Drug). Group comparisons, SalB versus Vehicle, Block1 *p=0.03, Block 2 *p=0.02. Right, shuttles per WS for Block 1 of Tests in FB-Devalued condition. Paired t-test: $t_{12}$=3.14 **p=0.009. **j**, Left, latency to first AR for Vehicle versus SalB Tests for FB-Devalued condition. Two-way ANOVA followed by Šídák's test: $F_{4,96}$=1.43 (Block x Drug); $F_{1,24}$=20.65 ***p=0.0001 (Drug). Group comparisons, SalB versus Vehicle, Block 1 *p=0.02, Block 2 *p=0.03. Right, latency to first AR during Block 1 of Tests for FB-Devalued condition. Paired t-test: $t_{12}$=3.35 **p=0.006. **k**, WS freezing for Block 1 of Vehicle versus SalB Tests in FB-Devalued condition. Paired t-test: $t_{12}$=4.61 ***p=0.0006. **l-o**, Diagrammatic summary of results. All data in the figure are shown as means. Error bars (s.e.m.) are shown for between-subject analyses.

Fig. 5

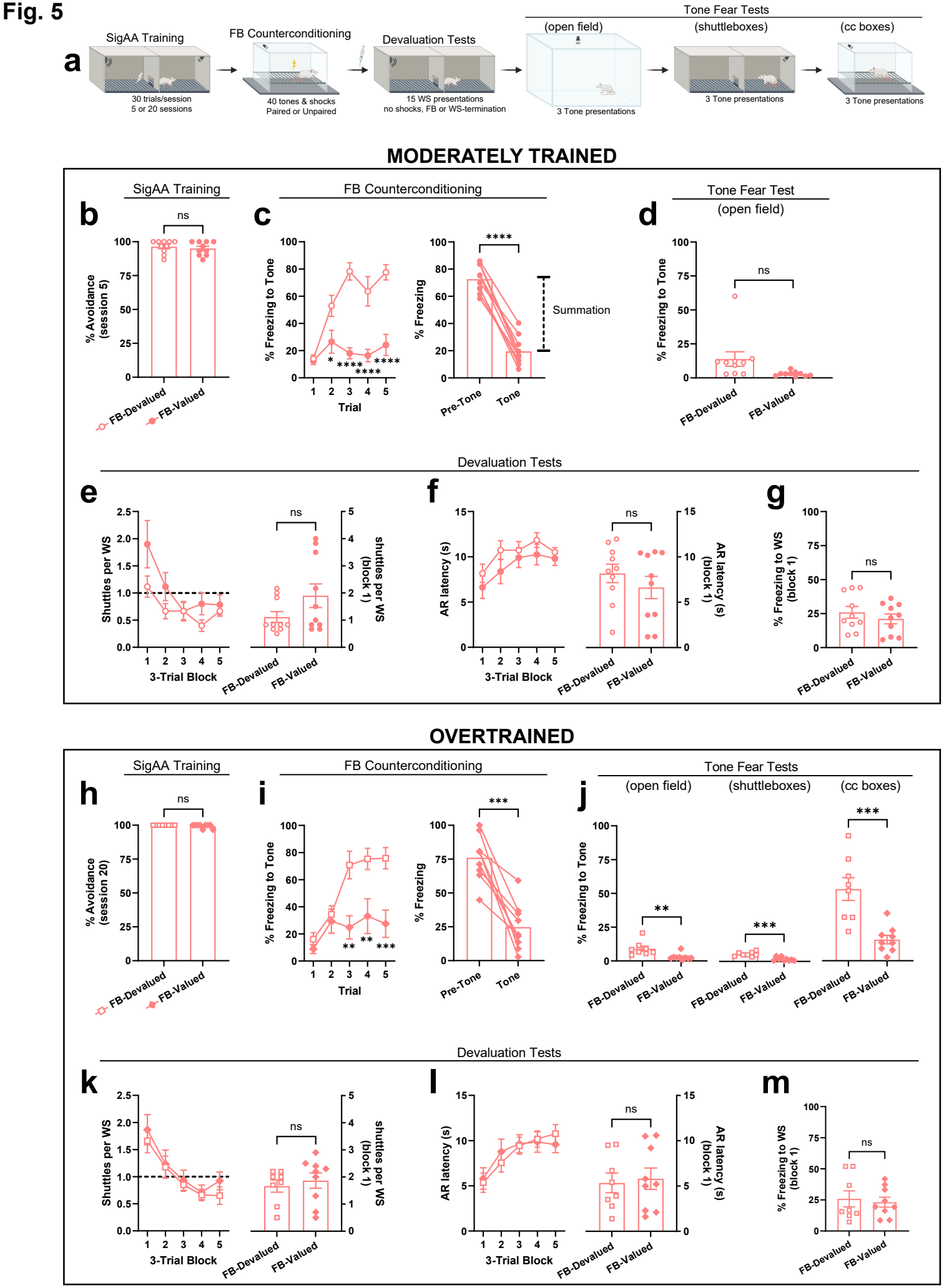

a
SigAA Training
30 trials/session
5 or 20 sessions
FB Counterconditioning
40 tones & shocks
Paired or Unpaired
Devaluation Tests
15 WS presentations
no shocks, FB or WS-termination
Tone Fear Tests
(open field)
3 Tone presentations
(shuttleboxes)
3 Tone presentations
(cc boxes)
3 Tone presentations
MODERATELY TRAINED
SigAA Training
b
% Avoidance (session 5)
ns
FB-Devalued
FB-Valued
FB Counterconditioning
c
% Freezing to Tone
Trial
% Freezing
Pre-Tone
Tone
Summation
Tone Fear Test
(open field)
d
% Freezing to Tone
ns
Devaluation Tests
e
Shuttles per WS
3-Trial Block
shuttles per WS (block 1)
f
AR latency (s)
AR latency (s) (block 1)
g
% Freezing to WS (block 1)
OVERTRAINED
SigAA Training
h
% Avoidance (session 20)
FB Counterconditioning
i
% Freezing to Tone
% Freezing
Tone Fear Tests
(open field)
(shuttleboxes)
(cc boxes)
j
Devaluation Tests
k
l
m

**Fig. 5 | Devaluation of response-produced safety signals after moderate training versus overtraining in female rats. a**, Schematic of the devaluation paradigm. Panels **b-g**: moderately trained rats. Panels **h-m**: overtrained rats. Note that moderately trained rats received only one Tone Fear Test in the open field. Additional Tone Fear Tests in the counterconditioning boxes and shuttleboxes were later added for overtrained rats to address differences between males and females. **b**, SigAA performance during session 5 of training for rats assigned to FB-Devalued (n=10) and FB-Valued (n=10) groups. Unpaired t-test: $t_{18}$=0.61 p=0.55. **c**, Left, tone freezing during the first 5 trials of counterconditioning. Two-way ANOVA followed by Šídák's test: $F_{4,72}$=9.183. Group comparisons, FB-Valued versus FB-Devalued, Trial 2 *p=0.03, Trial 3 ****p<0.0001, Trial 4 ****p<0.0001, Trial 5 ****p<0.0001. Right, Pre-Tone versus Tone freezing during first 5 trials of counterconditioning for FB-Valued rats. Paired t-test: $t_9$=17.11 ****p<0.0001. **d**, Tone freezing in the novel open field. Unpaired t-test: $t_{18}$=2.06 p=0.054. **e**, Right, shuttles per WS for both Devaluation Tests. Two-way ANOVA: $F_{4,72}$=1.77. Right, shuttles per WS for Block 1 of Devaluation Tests. Unpaired t-test: $t_{18}$=1.64 p=0.12. **f**, Left, latency to first AR for both Devaluation Tests. Two-way ANOVA: $F_{4,72}$=0.38. Right, latency to first AR during Block 1 of Devaluation Tests. Unpaired t-test: $t_{18}$=0.97, p=0.345. **g**, WS freezing for Block 1 of Devaluation Tests. Unpaired t-test: $t_{18}$=0.85, p=0.41. **h**, SigAA performance during session 20 of overtraining for rats assigned to FB-Devalued (n=8) and FB-Valued (n=9) groups. Unpaired t-test: $t_{15}$=1.42 p=0.18. **i**, Left, Tone freezing during the first 5 trials of counterconditioning. Two-way ANOVA followed by Šídák's test: $F_{4,60}$=3.85 **p=0.008. Group comparisons, FB-Valued versus FB-Devalued, Trial 3 **p=0.002, Trial 4 **p=0.005, Trial 5 ***p=0.0009. Right, Pre-Tone versus Tone freezing during first 5 trials of counterconditioning for FB-Valued rats. Paired t-test: $t_8$=6.39 ***p=0.0002. **j**, Tone freezing in the novel open field (left), shuttleboxes (middle), and counterconditioning boxes (right). Unpaired t-tests: $t_{15}$=3.20 **p=0.006 (left); $t_{15}$=5.21 ***p=0.0001 (middle); $t_{15}$=4.33 ***p=0.0006 (right). **k**, Left, shuttles per WS for both Devaluation Tests. Two-way ANOVA: $F_{4,60}$=0.33. Right, shuttles per WS for Block 1 of Devaluation Tests. Unpaired t-test: $t_{15}$=0.56 p=0.58. **l**, Left, latency to first AR for both Devaluation Tests. Two-way ANOVA: $F_{4,60}$=0.81. Right, latency to first AR during Block 1 of Devaluation Tests. Unpaired t-test: $t_{15}$=0.29 p=0.78. **m**, WS freezing for Block 1 of Devaluation Tests. Unpaired t-test: $t_{15}$=0.36 p=0.73. Data in **c** (right) and **i** (right) are shown as mean + individuals. All other data in the figure are shown as mean ± s.e.m.

**Fig. 6**

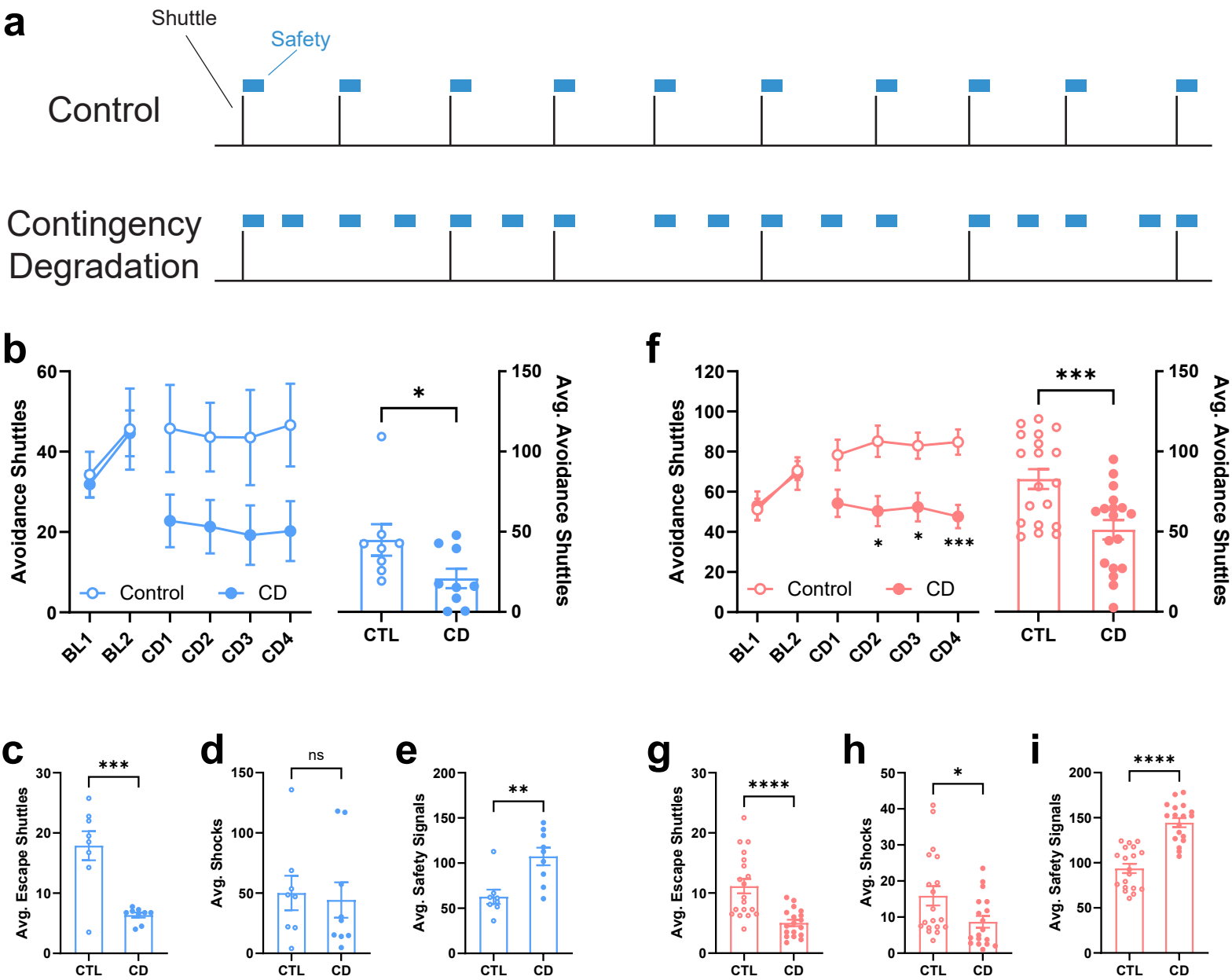
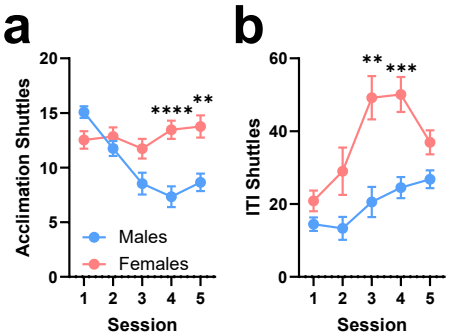

a
Shuttle
Safety
Control
Contingency Degradation
b
Avoidance Shuttles
Control
CD
BL1
BL2
CD1
CD2
CD3
CD4
CTL
CD
Avg. Avoidance Shuttles
f
c
Avg. Escape Shuttles
d
Avg. Shocks
ns
e
Avg. Safety Signals
g
h
i

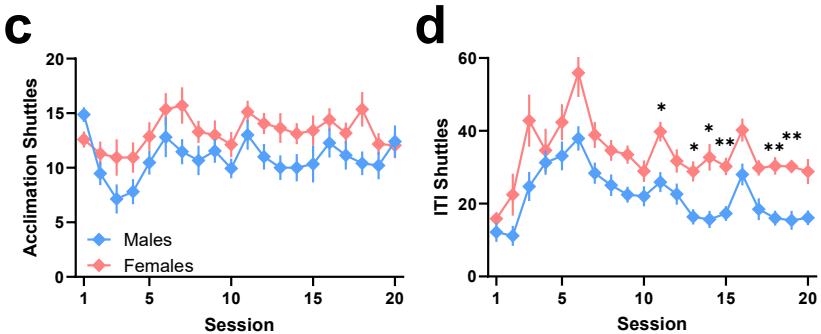
**Fig. 6 | Contingency degradation impairs goal-directed avoidance in male and female rats. a**, Schematic of the contingency degradation protocol. **b**, Left, Number of avoidance shuttles during the last two training sessions (BL1-2) and four contingency degradation sessions (CD1-4) for male rats tested at TAMU. Rats assigned to Control (n=8) and CD (n=9) groups performed similarly during the baseline phase. Two-way ANOVA: $F_{1,15}$=0.02 p=0.89. CD rats avoided less than Control rats during the contingency degradation phase. Two-way ANOVA: $F_{3,45}$=0.09 (Session x Group), $F_{1,15}$=4.57 *p=0.049. Right, avoidance shuttles per CD session. Unpaired t-test: $t_{15}$=2.14 *p=0.049. **c**, Average escape shuttles per CD session. Unpaired t-test: $t_{15}$=4.98 ***p=0.0002. **d**, Average shocks delivered per CD session. Unpaired t-test: $t_{15}$=0.28 p=0.78. **e**, Average safety signals per session. Unpaired t-test: $t_{15}$=3.52 **p=0.003. **f**, Left, Number of avoidance shuttles during the last two training sessions (BL1-2) and four contingency degradation sessions (CD1-4) for female rats tested at TAMU and NKI (combined here, see Extended Fig.6 for separate results). Rats assigned to Control (n=19) and CD (n=18) groups performed similarly during the baseline phase. Two-way ANOVA: $F_{1,35}$=0.15. CD rats avoided less than Control rats during the contingency degradation phase. Two-way ANOVA followed by Šídák's test: $F_{3,105}$=1.07 (Session x Group), $F_{1,35}$=13.34 ***p=0.0008 (Group). Group comparisons, CD versus Control, CD3 *p=0.01, CD4 *p=0.01, CD5 ***p=0.0005. Right,Average avoidance shuttles per CD session. Unpaired t-test: $t_{35}$=3.65 ***p=0.0008. **g**, Average escape shuttles per CD session. Unpaired t-test: $t_{35}$=4.54 ****p<0.0001. **h**, Average shocks per CD session. Unpaired t-test: $t_{35}$=2.28 *p=0.03. **i**, Average safety signals per CD session. Unpaired t-test: $t_{35}$=7.01 ****p<0.0001. For graphs showing individuals, circles represent subjects tested at TAMU and squares represent subjects tested at NKI. All data in the figure are shown as mean ± s.e.m.

# Ext. Fig. 1

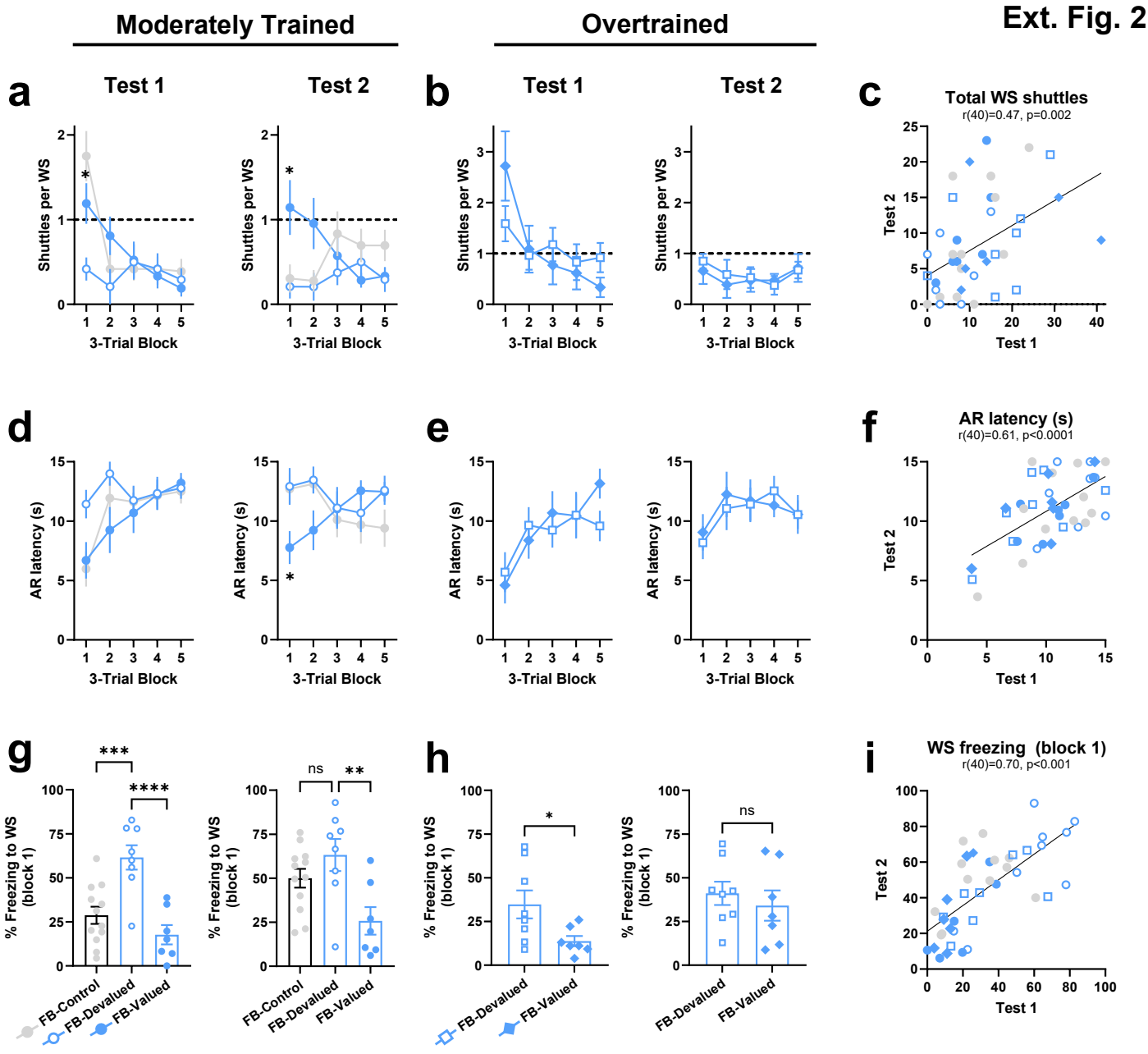
## Moderately Trained




## Overtrained

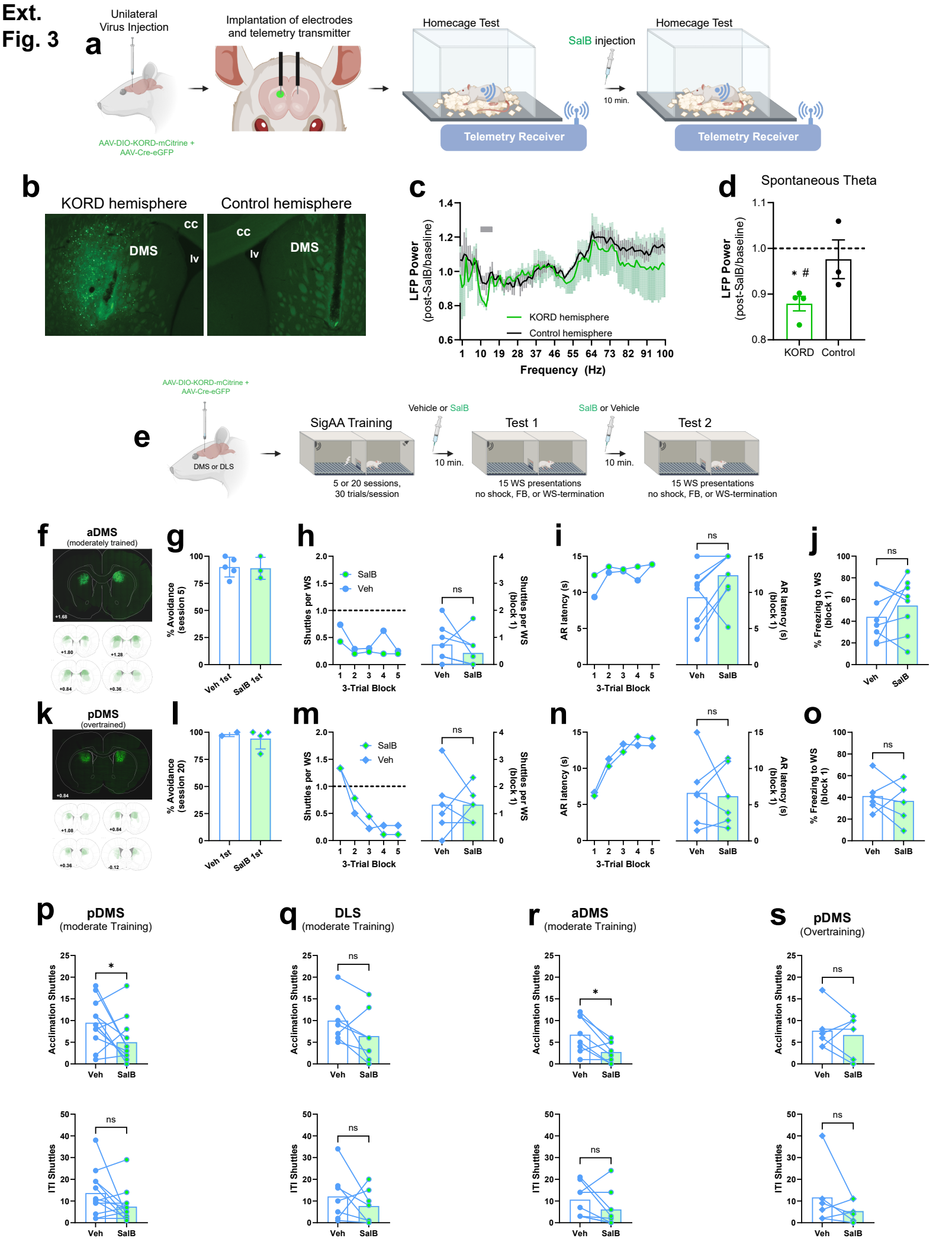


**Extended Fig. 1 | Female rats are slightly more active during SigAA training.** Shuttles during acclimation and ITI periods for rats depicted in Fig 1. All data were analyzed by two-way ANOVAs followed by Šídák's multiple comparisons tests. **a**, Acclimation shuttles for moderately trained male (n=35) and female (n=26) rats. Two-way ANOVA followed by Šídák's test: $F_{4,236}$=12.34. Group comparisons, Male versus Female, Session 4 ****p<0.0001, Session 5 **p= 0.0013. **b**, ITI shuttles for the same moderately trained rats. Two-way ANOVA followed by Šídák's test: $F_{4,236}$=3.76. Group comparisons, Male versus Female, Session 3 **p=0.0013, Session 4 ***p=0.0002. **c**, Acclimation shuttles for overtrained male (n=15) and female (n=17) rats. Two-way ANOVA: $F_{19,570}$=1.652. **d**, ITI shuttles for the same overtrained rats. Two-way ANOVA followed by Šídák's test: $F_{19,570}$=0.8593 (Session x Sex), $F_{1,30}$=20.56 (Sex). Group comparisons, Male versus Female, Session 11 *p=0.025, Session 13 *p= 0.025, Session 14 *p= 0.01, Session 15 **p= 0.006, Session 18 **p= 0.0019, Session 19 **p=0.0013. All data in the figure are shown as mean ± s.e.m.


Ext. Fig. 2
Moderately Trained
Overtrained
a
Test 1
Test 2
Shuttles per WS
3-Trial Block
b
Test 1
Test 2
Shuttles per WS
3-Trial Block
c
Total WS shuttles
r(40)=0.47, p=0.002
Test 2
Test 1
d
AR latency (s)
3-Trial Block
e
AR latency (s)
3-Trial Block
f
AR latency (s)
r(40)=0.61, p<0.0001
Test 2
Test 1
g
% Freezing to WS (block 1)
FB-Control
FB-Devalued
FB-Valued
h
% Freezing to WS (block 1)
FB-Devalued
FB-Valued
i
WS freezing (block 1)
r(40)=0.70, p<0.001
Test 2
Test 1

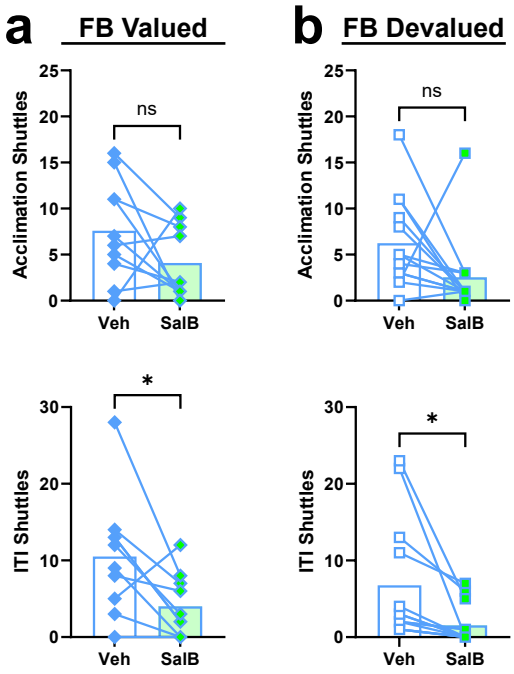
**Extended Fig. 2 | Devaluation test-retest reliability for male rats.** We conducted two Devaluation Tests in anticipation of within-subject studies exploring the neural substrates of goal-directed vs. habitual avoidance using DREADDs. These studies require that each rat receive a test after injection with Vehicle or the DREADD ligand (counterbalanced). For this strategy to be effective, it is important that there is high test-retest reliability for our avoidance and freezing measures. Here we show the results of Test 1 and Test 2 for all moderately trained and overtrained rats in Fig. 2. Importantly, correlations between Test 1 and Test 2 behavior were very strong for all three measures indicating high test-retest reliability. **a**, Shuttles per WS for moderately trained rats depicted in Fig. 2 during Devaluations Tests 1 (left) and 2 (right). Two-way ANOVAs followed by Dunnett's tests: Test 1, $F_{8,96}$=4.38; Test 2: $F_{8,96}$=4.595. Group comparisons, FB-Devalued versus FB-Valued, Test 1, Block 1 *p=0.034, Test 2, Block 1 *p=0.0495. **b**, Shuttles per WS for overtrained rats during Devaluation Tests 1 (left) and 2 (right). Two-way ANOVAs: Test 1, $F_{4,52}$=3.02; Test 2, $F_{4,52}$=0.3925. **c**, Pearson correlation of total WS shuttles (Test 1 vs Test 2) for all rats depicted in **a** & **b** ($r_{40}$=0.4680, p=0.0018). **d**, Latency to first AR for moderately trained rats during Devaluation Tests 1 (left) and 2 (right). Two-way ANOVAs followed by Dunnett's tests: Test 1, $F_{8,96}$=1.82; Test 2, $F_{8,96}$=2.85. Group comparisons, FB-Devalued versus FB-Valued, Test 2, Block1 *p=0.05. **e**, Latency to first AR for overtrained rats during Devaluation Tests 1 (left) and 2 (right). Two-way ANOVAs: Test 1, $F_{4,52}$=1.43; Test 2, $F_{4,52}$=0.35. **f**, Pearson correlation of average latency to first AR (Test 1 vs Test 2) for all rats depicted in **d** & **e** ($r_{40}$=0.6063, p<0.0001). **g**, WS freezing during Block 1 for moderately trained rats in Devaluation Tests 1 (left) and 2 (right). One-way ANOVAs followed by Dunnett's tests: Test 1, $F_{2,24}$=13.87; Test 2, $F_{2,24}$=5.85. Group comparisons, FB-Devalued versus FB-Control, Test 1 ***p=0.0006, FB-Devalued versus FB-Valued, Test 1 ****p<0.0001, Test 2 **p=0.0047. **h**, WS freezing during Block 1 for overtrained rats in Devaluation Tests 1 (left) and 2 (right). Unpaired t-tests: Test 1, $t_{13}$=2.34 *p=0.036; Test 2, $t_{13}$=0.65 p=0.52. **i**, Pearson correlation of average WS freezing (Block 1; Test 1 vs Test 2) for all rats depicted in **g** & **h** ($r_{40}$=0.7031, p<0.0001). Data in panels **a**, **b**, **d**, **e**, **g** & **h** are shown as mean ± s.e.m.

Ext. Fig. 3


a
Unilateral Virus Injection
AAV-DIO-KORD-mCitrine + AAV-Cre-eGFP
Implantation of electrodes and telemetry transmitter
Homecage Test
Telemetry Receiver
SalB injection
10 min.
Homecage Test
Telemetry Receiver
b
KORD hemisphere
Control hemisphere
cc
DMS
lv
c
LFP Power (post-SalB/baseline)
KORD hemisphere
Control hemisphere
Frequency (Hz)
d
Spontaneous Theta
LFP Power (post-SalB/baseline)
KORD
Control
e
AAV-DIO-KORD-mCitrine + AAV-Cre-eGFP
DMS or DLS
SigAA Training
5 or 20 sessions, 30 trials/session
Vehicle or SalB
10 min.
Test 1
15 WS presentations no shock, FB, or WS-termination
SalB or Vehicle
10 min.
Test 2
15 WS presentations no shock, FB, or WS-termination
f
aDMS
(moderately trained)
g
% Avoidance (session 5)
Veh 1st
SalB 1st
h
SalB
Veh
Shuttles per WS
3-Trial Block
ns
Shuttles per WS (block 1)
i
AR latency (s)
3-Trial Block
ns
AR latency (s) (block 1)
j
% Freezing to WS (block 1)
ns
k
pDMS
(overtrained)
l
% Avoidance (session 20)
Veh 1st
SalB 1st
m
SalB
Veh
Shuttles per WS
3-Trial Block
ns
Shuttles per WS (block 1)
n
AR latency (s)
3-Trial Block
ns
AR latency (s) (block 1)
o
% Freezing to WS (block 1)
ns
p
pDMS
(moderate Training)
Acclimation Shuttles
*
ITI Shuttles
ns
q
DLS
(moderate Training)
Acclimation Shuttles
ns
ITI Shuttles
ns
r
aDMS
(moderate Training)
Acclimation Shuttles
*
ITI Shuttles
ns
s
pDMS
(Overtraining)
Acclimation Shuttles
ns
ITI Shuttles
ns
Veh
SalB

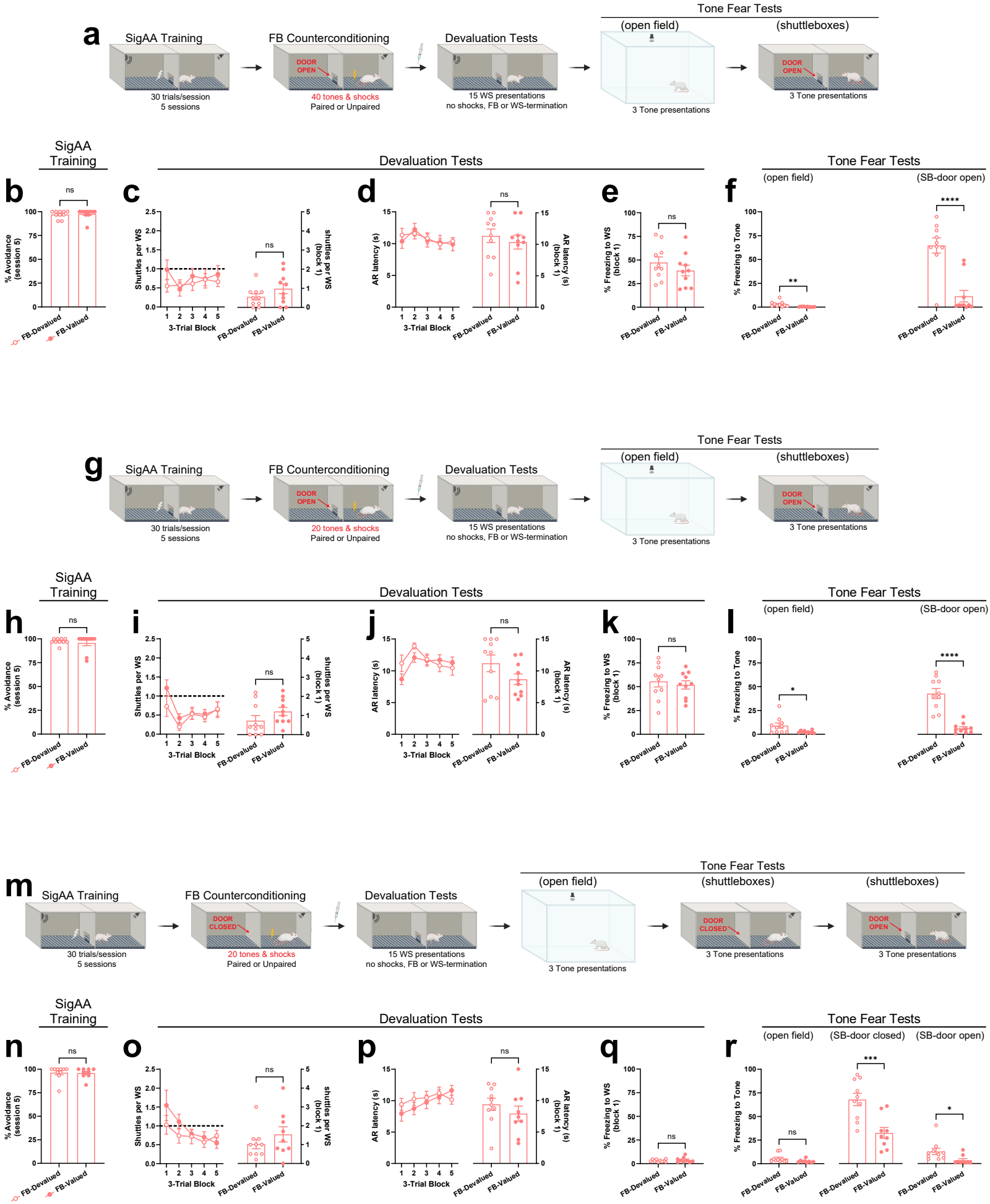
**Extended Fig. 3 | KORD validation, relevant null effects, and non-avoidance shuttling. a**, Schematic of *in vivo* electrophysiology experiment to validate KORD/SalB inhibition in DMS. **b**, Example of unilateral KORD expression in DMS and bilateral electrode tracks. **c**, Spontaneous LFP power in KORD versus Control hemispheres post-SalB injection (normalized to baseline, n=4 rats). Gray Bar denotes theta band. **d**, Spontaneous theta band LFP power in KORD versus Control hemispheres post-SalB injection (normalized to baseline). Paired t-tests: baseline versus post-SalB theta in KORD hemisphere, $t_3$=7.67 #p<0.01; post-SalB theta in KORD versus Control hemisphere, $t_2$=2.81 *p=0.05; baseline versus post-SalB theta in Control hemisphere, $t_2$=0.56 p=0.32. **e**, Schematic of the experimental paradigm evaluating aDMS after moderate training (n=8) and pDMS after overtraining (n=6). **f**, Top, example of bilateral KORD expression in aDMS. Bottom, Semitransparent overlays illustrating KORD expression in aDMS for individuals across the A-P axis. A-P coordinates are in millimeters relative to Bregma. **g**, SigAA performance during session 5 of training for aDMS rats assigned to receive Vehicle versus SalB in Test 1. Unpaired t-test: $t_6$=0.16 p=0.88. **h**, Left, shuttles per WS for Vehicle versus SalB Tests in aDMS rats. Two-way ANOVA: $F_{4,56}$=0.75. Right, shuttles per WS for Block 1 of Tests. Paired t-test: $t_7$=1.195 p=0.27. **i**, Left, latency to first AR for Vehicle versus SalB Tests in aDMS rats. Two-way ANOVA: $F_{4,56}$=0.86. Right, latency to first AR during Block 1 of Tests. Paired t-test: $t_7$=2.11 p=0.07. **j**, WS freezing for Block 1 of Vehicle versus SalB Tests in aDMS rats. Paired t-test: $t_7$=1.08 p=0.32. **k**, Top, example of bilateral KORD expression in pDMS. Bottom, semitransparent overlays illustrating KORD expression in pDMS for individuals across the A-P axis. A-P coordinates are in millimeters relative to Bregma. **l**, SigAA performance during session 20 of training for pDMS rats assigned to receive Vehicle versus SalB in Test 1. Unpaired t-test: $t_4$=0.58 p=0.60. **m**, Left, shuttles per WS for Vehicle versus SalB Tests in overtrained pDMS rats. Two-way ANOVA: $F_{4,40}$=0.48. Right, shuttles per WS for Block 1 of Tests. Paired t-test: $t_5$=0.00 p>0.99. **n**, Left, latency to first AR for Vehicle versus SalB Tests in overtrained pDMS rats. Two-way ANOVA: $F_{4,40}$=0.415. Right, latency to first AR during Block 1 of Tests. Paired t-test: $t_5$=0.24 p=0.82. **o**, WS freezing for Block 1 of Vehicle versus SalB Tests in overtrained pDMS rats. Paired t-test: $t_5$=0.575 p=0.59. **p**, Top, Acclimation shuttles for Vehicle versus SalB Tests in moderately trained pDMS rats. Paired t-test: $t_{11}$=2.305 *p=0.04. Bottom, ITI shuttles for the same rats. Paired t-test: $t_{11}$=1.84 p=0.09. **q**, Acclimation shuttles for Vehicle versus SalB Tests in moderately trained DLS rats. Paired t-test: $t_6$=1.77 p=0.13. Bottom, ITI shuttles for the same rats. Paired t-test: $t_6$=0.84 p=0.43. **r**, Acclimation shuttles for Vehicle versus SalB Tests in moderately trained aDMS rats. Paired t-test: $t_7$=3.35 *p=0.01. Bottom, ITI shuttles for the same rats. Paired t-test: $t_7$=1.48 p=0.18. **s**, Acclimation shuttles for Vehicle versus SalB Tests in overtrained pDMS rats. Paired t-test: $t_5$=0.51 p=0.63. Bottom, ITI shuttles for the same rats. Paired t-test: $t_5$=1.26 p=0.265. cc: corpus callosum, lv: lateral ventricle. All data in the figure are shown as means. Error bars (s.e.m.) are shown for between-subject analyses.

# Ext. Fig. 4


a
FB Valued
b
FB Devalued
Acclimation Shuttles
ITI Shuttles
ns
*
Veh
SalB
0
5
10
15
20
25
30

**Extended Fig. 4 | Acclimation and ITI shuttles for DLS habit experiment. a**, top, Acclimation shuttles for Vehicle versus SalB Tests in overtrained, FB-Valued rats depicted in Fig. 4. Paired t-test: $t_9$=1.62 p=0.14. Bottom, ITI shuttles for the same rats. Paired t-test: $t_9$=2.78 *p=0.02. **b**, Acclimation shuttles for Vehicle versus SalB Tests in overtrained, FB-Devalued rats (n=13). Paired t-test: $t_{12}$=1.87 p=0.09. Bottom, ITI shuttles for the same rats. Paired t-test: $t_{12}$=2.76 *p=0.02. All data in the figure are shown as mean + individual scores.

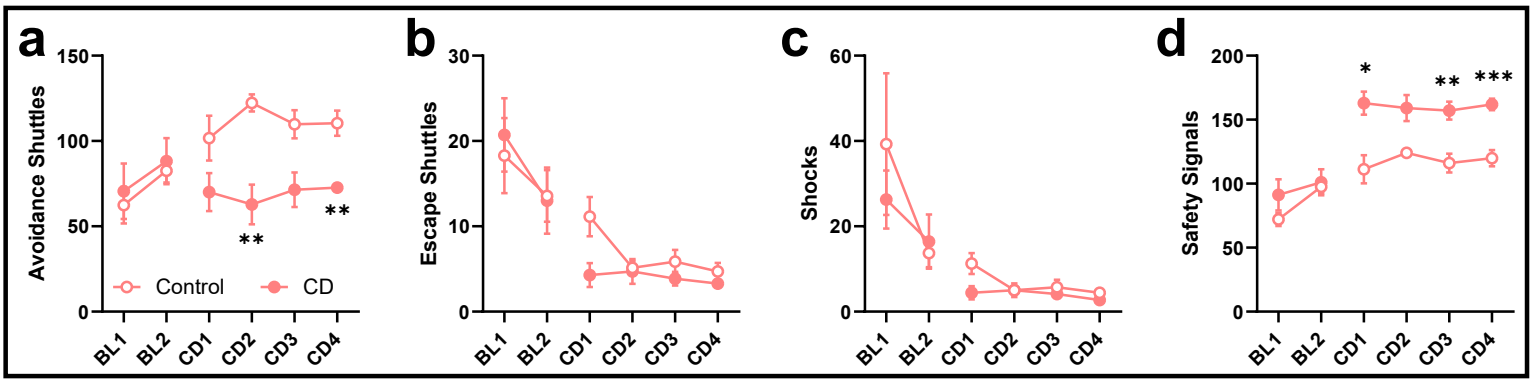
# Ext. Fig. 5


a
SigAA Training
FB Counterconditioning
Devaluation Tests
Tone Fear Tests
(open field)
(shuttleboxes)
DOOR OPEN
30 trials/session
5 sessions
40 tones & shocks
Paired or Unpaired
15 WS presentations
no shocks, FB or WS-termination
3 Tone presentations
b
c
d
e
f
(SB-door open)
% Avoidance (session 5)
Shuttles per WS
shuttles per WS (block 1)
AR latency (s)
AR latency (s) (block 1)
% Freezing to WS (block 1)
% Freezing to Tone
3-Trial Block
FB-Devalued
FB-Valued
ns
**
****
g
20 tones & shocks
h
i
j
k
l
*
m
DOOR CLOSED
(shuttleboxes)
n
o
p
q
r
(SB-door closed)
***

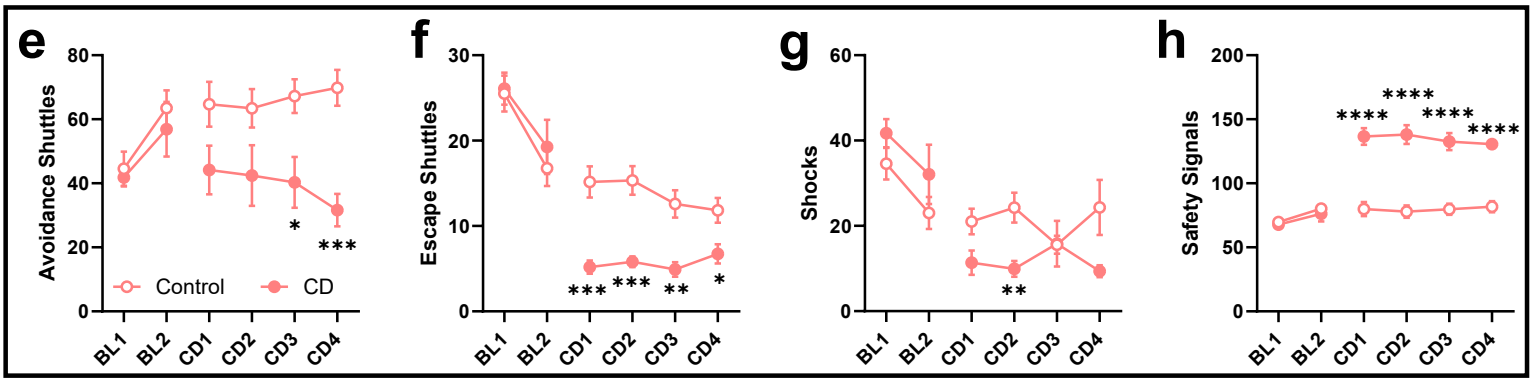
**Extended Fig. 5 | Devaluation of response-produced safety signals in the SigAA context (females).** **a**, Schematic of the OPEN40 experimental protocol. Counterconditioning occurred in the shuttleboxes and was otherwise identical to previous experiments. **b**, SigAA performance during session 5 of training for rats assigned to FB-Devalued (n=10) and FB-Valued (n=10) groups. Unpaired t-test: $t_{18}$=0.32 p=0.75. **c**, Left, shuttles per WS for both Devaluation Tests. Two-way ANOVA: $F_{4,72}$=1.19. Right, shuttles per WS for Block 1 of Devaluation Tests. Unpaired t-test: $t_{18}$=1.42 p=0.17. **d**, Left, latency to first AR for both Devaluation Tests. Two-way ANOVA: $F_{4,72}$=0.41. Right, latency to first AR during Block 1 of Devaluation Tests. Unpaired t-test: $t_{18}$=0.64 p=0.53. **e**, WS freezing for Block 1 of Devaluation Tests. Unpaired t-test: $t_{18}$=1.06 p=0.30. **f**, Tone freezing in the novel open field (left) and shuttleboxes (door open; right). Unpaired t-tests: $t_{18}$=2.997 **p=0.008 (left), $t_{18}$=5.34 ****p<0.0001 (right). **g**, Schematic of the OPEN20 experimental protocol, which was identical to OPEN40 except only 20 tones & shocks were delivered during counterconditioning. **h**, SigAA performance during session 5 of training for rats assigned to FB-Devalued (n=10) and FB-Valued (n=10) groups. Unpaired t-test: $t_{18}$=0.55 p=0.59. **i**, Left, shuttles per WS for both Devaluation Tests. Two-way ANOVA: $F_{4,72}$=1.265. Right, shuttles per WS for Block 1 of Devaluation Tests. Unpaired t-test: $t_{18}$=1.405 p=0.18. **j**, Left, latency to first AR for both Devaluation Tests. Two-way ANOVA: $F_{4,72}$=2.25. Right, Average latency to first AR during Block 1 of Devaluation Tests. Unpaired t-test: $t_{18}$=1.6 p=0.12. **k**, WS freezing for Block 1 of Devaluation Tests. Unpaired t-test: $t_{18}$=0.51 p=0.62. **l**, Tone freezing in the novel open field (left) and shuttleboxes (door open; right). Unpaired t-tests: $t_{18}$=2.42 *p=0.03 (left), $t_{18}$=6.64 ****p<0.0001 (right). **m**, Schematic of the CLOSED20 experimental protocol, which was identical to OPEN20 except the door connecting shuttlebox chamber sides was closed during counterconditioning and an additional Tone Fear Test was conducted with the shuttlebox door closed. **n**, SigAA performance during session 5 of training for rats assigned to FB-Devalued (n=10) and FB-Valued (n=9) groups. Unpaired t-test: $t_{17}$=0.14 p=0.89. **o**, Left, shuttles per WS for both Devaluation Tests. Two-way ANOVA: $F_{4,68}$=1.49. WS shuttles for Block 1 of Devaluation Tests. Unpaired t-test: $t_{17}$=1.11 p=0.28. **p**, Left, latency to first AR for both Devaluation Tests. Two-way ANOVA: $F_{4,68}$=1.32. Right, latency to first AR during Block 1 of Devaluation Tests. Unpaired t-test: $t_{17}$=0.96 p=0.35. **q**, WS freezing for Block 1 of Devaluation Tests. Unpaired t-test: $t_{17}$=0.84 p=0.41. **r**, Tone freezing in the novel open field (left), shuttleboxes with door closed (middle), and shuttleboxes with door open (right). Unpaired t-tests: $t_{17}$=1.79 p=0.09 (left), $t_{17}$=4.00 ***p=0.0009 (middle), $t_{17}$=2.38 *p=0.03 (right). All data in the figure are shown as mean ± s.e.m.

**Ext. Fig. 6**

## NKI




## TAMU



**Extended Fig. 6 | Replication of contingency degradation effect on goal-directed avoidance in females at separate institutions.** Identical experiments were conducted in parallel at NKI (panels **a**-**d**) and TAMU (panels **e**-**h**) in females. Results are presented separately here but were combined for analysis in Fig. 6 given a similar pattern of results. Note that rats were matched into equal groups before the CD phase and there were no differences in baseline behavior (BL1-2) for any measure in either experiment. **a**, Right, CD rats (n=7) avoided less than Control rats (n=7) during the contingency degradation phase. Two-way ANOVA followed by Šídák's test. $F_{3,36}$=1.13 (Session x Group), $F_{1,12}$=23.06 (Group). Group comparisons, CD versus Control, CD2 **p=0.006, CD4 **p=0.0060. **b**, Right, CD rats escaped less than Control rats during the contingency degradation phase. Two-way ANOVA: $F_{3,36}$=2.71 (Session x Group), $F_{1,12}$=6.00 *p=0.03 (Group). **c**, Right, CD and Control rats received similar patterns of shock during the contingency degradation phase. Two-way ANOVA: $F_{3,36}$=2.36 (Session x Group), $F_{1,12}$=4.13 (Group). **d**, Right, CD rats received more safety signals than Control rats during the contingency degradation phase. Two-way ANOVA followed by Šídák's test: $F_{3,36}$=0.48 (Session x Group), $F_{1,12}$=37.68 (Group). Group comparisons, CD versus Control, CD1 *p=0.015, CD3 **p=0.007, CD4 ***p=0.0009. **e**, Right, CD rats (n=11) avoided less than Control rats (n=12) during the contingency degradation phase. Two-way ANOVA followed by Šídák's test: $F_{3,63}$=2.18 (Session x Group), $F_{1,21}$=10.10 (Group). Group comparisons, CD verus Control, CD3 *p= 0.045, CD4 ***p=0.0002. **f**, Right, CD rats escaped less than Control rats during the contingency degradation phase. Two-way ANOVA followed by Šídák's test: $F_{3,63}$=3.423. Group comparisons, CD versus Control, CD1 ***p=0.0006, CD2 ***p=0.0005, CD3 **p=0.002, CD4 *p=0.045). **g**, Right, CD rats received fewer shocks than Control rats during the contingency degradation phase. Two-way ANOVA followed by Šídák's tests: $F_{3,63}$=3.12. Group comparisons, CD versus Control, CD2 **p=0.009. **h**, Right, CD rats received more safety signals than Control rats during the contingency degradation phase. Two-way ANOVA followed by Šídák's test: $F_{3,63}$=1.23 (Session x Group), $F_{1,21}$=63.09. Group comparisons, CD versus Control, CD1 ****p<0.0001, CD2 ****p<0.0001, CD3 ****p<0.0001, ****p<0.0001. BL: baseline, CD: contingency degradation, TAMU: Texas A&M University, NKI: Nathan Kline Institute. All data in the figure are shown as mean ± s.e.m.