1 **Sex differences in fear regulation and reward seeking behaviors in a fear-safety-reward**

2 **discrimination task**

3

4

5 Abbreviated title: Sex differences in conditioned safety

6

7

8

9 Eliza M. Greiner[1], Makenzie R. Norris[1], Iris Müller [1,2], Ka H. Ng[1,2], Susan Sangha[1,2]*

10

11 1. Department of Psychological Sciences

12 2. Purdue Institute for Integrative Neuroscience

13 Purdue University, West Lafayette, IN, USA 47907

14

15 *Correspondence to:

16 Susan Sangha

17 Department of Psychological Sciences

18 Purdue University

19 703 Third St., West Lafayette, IN 47907

20 (t) 765.494.0648

21 (e) sangha@purdue.edu

22

23

24 Number of figures: 5

25 Number of tables: 1

26

27

Abstract

Reward availability and the potential for danger or safety potently regulates emotion. Despite women being more likely than men to develop emotion dysregulation disorders, there are few studies investigating fear, safety and reward regulation in females. Here, we show that female Long Evans rats do not suppress conditioned freezing in the presence of a safety cue, nor do they extinguish their freezing response, whereas males do both. Females did show evidence of conditioned inhibition of darting, however. Females were also more reward responsive during the reward cue until the first footshock exposure, at which point there were no sex differences in reward seeking to the reward cue. In summary, females showed a significantly different behavioral profile than males in a task that tests the ability to discriminate among fear, safety and reward cues. This paradigm offers a great opportunity to test for mechanisms that are generating these behavioral sex differences.

Introduction

Clinical disorders arising from maladaptive emotion regulation present a large burden on society worldwide. Many of these disorders show comorbidity, for example, addiction with anxiety disorders (Grant et al. 2016). Cues predicting something aversive elicit avoidance and fear behaviors whereas cues predicting reward elicit approach and reward-seeking behaviors. Cues signifying safety have the power to modulate fear and reward-seeking behaviors by informing the organism whether or not the environment is safe (Walasek et al. 1995). Thus, safety, fear and reward behaviors, and the circuitries governing these behaviors, are intertwined. The majority of studies on reward and fear processing have been conducted in parallel, investigating the circuitries separately in primarily male subjects. If we hope to understand and treat comorbid disorders resulting from maladaptive emotion regulation, increased efforts in investigating how these circuitries integrate their functions to influence behavior is needed in both male and female subjects.

Our laboratory has designed and validated a behavioral task in which fear, safety and reward cues are learned within the same session allowing us to assess the animal's ability to discriminate among these cues (Sangha et al. 2013; Sangha, Robinson, et al. 2014; Sangha, Greba, et al. 2014; Ng et al. 2018). Rats are exposed to cues associated with safety, fear (fear cue paired with footshock), and reward (reward cue paired with sucrose). Male rats consistently learn to discriminate among safety, fear and reward cues to 1) suppress conditioned freezing in the presence of a safety cue (fear+safety cue), and 2) increase reward seeking when reward is available (reward cue) (Sangha et al. 2013; Sangha, Robinson, et al. 2014; Sangha, Greba, et al. 2014; Ng et al. 2018). This paradigm also allows us to investigate how safety cues can regulate both fear and reward behaviors. Evidence suggests that reward learning mechanisms overlap with safety learning (Sangha et al. 2013; Rescorla 1969; Walasek et al. 1995; Rogan et al. 2005; Pollak et al. 2008; Tanimoto et al. 2004; Leknes et al. 2011). For example, learned safety can act as a behavioral antidepressant in mice (Pollak et al. 2008), and animals will perform certain behaviors in order to turn on a safety signal (Rescorla 1969; Rogan et al. 2005). Within the amygdala we have shown a subpopulation of neurons responding with the same level of excitation or inhibition during both the reward and safety cues (Sangha et al. 2013). Additionally, we have shown inactivation of the prelimbic or infralimbic cortices of the ventromedial prefrontal cortex have differential effects on reward and safety discrimination, respectively (Sangha, Robinson, et al. 2014). Thus, in male rats, our lab has already shown a

73   critical involvement of the corticoamygdalar circuit in learning this fear-safety-reward cue

74   discrimination.

75

76   Much of the research investigating fear regulation mechanisms have exclusively used male

77   subjects. But, women are more than twice as likely to develop Post-Traumatic Stress Disorder

78   (PTSD) than men, with females having a lifetime prevalence of 8.5% in contrast to 3.4% in

79   males (Mclean et al. 2011). In a study using male Vietnam veterans, PTSD patients show

80   impairments in suppressing their fear response in the presence of a safety cue (Jovanovic et al.

81   2009). In fear studies that have included female rats, it has been shown that females exhibit

82   lower levels of freezing behavior than male rats after repeated fear cue presentations (Daviu et

83   al. 2014). These findings have been thought to indicate a difficulty in fear conditioning learning

84   in female rats. A more recent experiment has identified that approximately 40% of female rats

85   tested exhibit an alternate fear behavior in the form of fast paced movements called 'darting';

86   this was only seen in approximately 10% of male rats tested (Gruene et al. 2015). In reward

87   studies that have included female rats, significant sex differences in response to drugs of abuse

88   have been seen, with female rats consistently self-administering drugs more rapidly than males

89   (Becker & Koob 2016). This seems at odds with the clinical evidence showing men typically

90   consume more drugs, such as alcohol, than women. However, data also indicate that alcohol

91   consumption for men and women are becoming increasingly similar (Keyes et al. 2008). And,

92   women who become dependent on opioid, cannabis or alcohol progress faster than men

93   through the stages of initial use to dependence (Hernandez-Avila et al. 2004), mimicking the

94   data reporting faster escalation of drug self-administration in female rats.

95

96   Taken together, we hypothesized there would be sex differences in the ability to express clear

97   discrimination among fear, safety and reward cues. The inability of male PTSD patients to learn

98   safety signaling has been labeled a biomarker of the disorder (Jovanovic et al. 2012). Due to

99   sex-related differences in human diagnosis of PTSD, with women diagnosed at rates twice that

100   of men (Glover et al. 2015), any differences female rats have in the learning or retention of

101   safety signals could steer towards further research on the neurological processes underlying

102   these variations.

103 **Results**

104

105 *Female rats spend more time reward seeking during reward pre-training*

106 All rats first underwent 5 reward pre-training sessions in which the reward cue was paired with

107 sucrose delivery. The percent time spent at or in the reward port during each reward cue across

108 each reward session was quantified (Figure 1b). Two-way repeated-measures ANOVAs showed

109 main effects of session ($F_{(4,136)}=5.395$, p=0.0005) and sex ($F_{(1,34)}=10.83$, p=.0023), but no

110 significant interaction ($F_{(4,136)}=0.9031$, p=0.4641). *Post hoc* Sidak's multiple comparisons test

111 showed females spent significantly more time reward seeking during the reward cue than males

112 for sessions R2 (p=0.0274), R3 (p=0.0151) and R5 (p=0.0041). The latency, in seconds, to

113 enter the port post-cue onset was calculated for each reward cue presentation across all

114 sessions (Figure 1c). Two-way repeated-measures ANOVAs showed a main effect of sex

115 ($F_{(1,34)}=20.37$, p<.0001), but no significant interaction ($F_{(4,136)}=1.684$, p=0.1571) or main

116 effect of session ($F_{(4,136)}=0.7755$, p=0.5429). *Post hoc* Sidak's multiple comparisons test

117 showed females were significantly faster to enter the port than males during the last 3 reward

118 sessions (R3, p=0.001; R4, p=0.0391; R5, p=0.0014). Taken together, female rats consistently

119 spent more time than males in the reward port during the reward cue in reward pre-training

120 sessions.

121

122 *Female rats do not show conditioned inhibition of freezing but instead conditioned inhibition of*

123 *darting*

124 After reward pre-training, rats were then exposed to sessions consisting of reward, fear and

125 safety cues. The reward cue and sucrose reward were the same as the reward pre-training

126 sessions. The fear cue was paired with a 0.5mA footshock, and both the safety cue and

127 fear+safety cue did not result in footshock or sucrose.

128

129 The percent time spent at or in the reward port during each cue across session was quantified

130 for each DC session (Figure 2b). Two-way repeated-measures ANOVAs showed a significant

131 cue by sex effect, as well as main effects of cue and sex for DC1 (Table 1). *Post hoc* Sidak's

132 multiple comparisons test showed that, during DC1, females spent significantly more time

133 reward seeking during the reward cue compared to males (p<0.001), consistent to what was

134 seen in reward pre-training. For the remaining DC2-4 sessions, a main effect of cue was

135 observed (Table 1) and *post hoc* Sidak's multiple comparisons test showed that both male and

136 female rats spent significantly more time reward seeking during the reward cue compared to all

5

137   other cues (p<0.0001), with no significant differences between the males and females. Thus,

138   the noticeable increase in reward seeking in the females, that was seen during reward pre-

139   training, dissipated by the 2nd DC session.

140

141   The percent time freezing during each cue across session was quantified for each DC session

142   (Figure 2c). Two-way repeated-measures ANOVAs showed a significant cue by sex effect for

143   sessions DC2-4, as well as main effects of cue and sex for every session (Table 1). *Post hoc*

144   Sidak's multiple comparisons tests showed that, for every session, females displayed

145   significantly more freezing to the fear+safety cue compared to males (DC1, p=0.0313; DC2,

146   p=0.007; DC3, p=0.0007; DC4, p<0.0001). Females also showed significantly higher freezing

147   levels to the fear cue compared to males during DC2 (p=0.0111). Males showed a significant

148   reduction in freezing levels to the fear+safety cue compared to the fear cue during sessions

149   DC3 (p=0.0156) and DC4 (p<0.0001), thus showing significant conditioned inhibition of freezing.

150   Females did not show a significant reduction for any session.

151

152   The number of darts during each cue was also quantified for each DC session and expressed

153   as a dart rate (Figure 2d; # darts/20s cue). Darting behavior during cue presentation was largely

154   absent until DC3 and DC4. Two-way repeated-measures ANOVAs showed a significant cue by

155   sex effect for DC4, as well as main effects of cue, for DC2-4, and sex, for DC1 and DC4 (Table

156   1). *Post hoc* Sidak's multiple comparisons test showed that, during DC4, females expressed

157   more darting behavior compared to males during both the fear cue (p<0.0001) and the

158   fear+safety cue (p=0.0079). The females also showed a significant reduction in darting behavior

159   during the fear+safety cue compared to the fear cue during DC4 (p=0.0166), suggesting some

160   level of conditioned inhibition present in the females, when darting is taken into account.

161

162   *Female rats do not show extinction of freezing*

163   The day after the last DC session all rats underwent fear and reward extinction within the same

164   session. That is, both the fear and reward cues were presented within the same training

165   session, without footshocks or sucrose presentations.

166

167   For reward extinction training, there was no main effect of reward trial ($F_{(19,646)}=1.526$,

168   p=0.0704) or sex ($F_{(1,34)}=1.31$, p=0.2603) and no interaction ($F_{(19,646)}=0.8927$, p=0.5924);

169   there was also no significant difference between male and female groups for any trial (Figure

170   3Bi). For the extinction test (Figure 3Bii), there was a main effect of cue (2-way RM ANOVA;

171    $F(3,102)=15.68$, $p<0.0001$) and sex ($F(1,34)=6.156$, $p=0.0182$). *Post hoc* Sidak's multiple

172   comparisons test showed females had significantly more port activity than males during the

173   safety cue ($p=0.0452$).

174

175   To assess fear extinction the averaged percent time freezing during each trial of fear extinction

176   training was calculated (Figure 3Ci). There was a main effect of fear trial (2-way RM ANOVA;

177    $F(19, 646)=7.69$, $p<0.0001$) and sex (2-way RM ANOVA; $F(1, 34)=4.607$, $p=0.0391$), but no

178   significant interaction ($F(19, 646)=1.566$, $p=0.059$). Compared to trial 1, males showed

179   significantly reduced freezing in extinction trials 8, 9 and 11-20 (*post hoc* Tukey's multiple

180   comparisons test, $p<0.05$), demonstrating good fear extinction. In contrast, females only

181   showed a significant reduction in freezing during trial 19 compared to the first trial (*post hoc*

182   Tukey's multiple comparisons test, $p=0.0394$), demonstrating relatively absent fear extinction.

183   For the extinction test (Figure 3Cii), there was a main effect of cue ($F(3,102)=134.7$, $p<0.0001$)

184   and sex ($F(1, 34)=6.217$, $p=0.0177$), as well as a significant interaction of cue X sex ($F(3,$

185    $102)=3.481$, $p=0.0187$). *Post hoc* Sidak's multiple comparisons test showed that females froze

186   significantly more than males to the fear ($p=0.0146$) and fear+safety ($p=0.0091$) cues.

187

188   We also assessed darting levels across extinction in response to each fear cue presentation

189   (Figure 3Di). During extinction training, there was a main effect of sex ($F(1,34)=4.816$,

190    $p=0.0351$), but no effect of trial ($F(19,646)=0.6941$, $p=0.8268$) and no significant interaction

191   ($F(19, 646)=1.083$, $p=0.3640$). *Post hoc* Sidak's multiple comparisons test showed no significant

192   differences between males and females for any trial. For the extinction test (Figure 3Dii), there

193   was a significant cue X sex interaction ($F(3,102)=4.447$, $p=0.0056$), as well as a main effect of

194   both cue ($F(3, 102)=4.248$, $p=0.0072$) and sex ($F1, 34)=4.834$, $p=0.0348$). Females showed a

195   significantly higher dart rate than males during the fear cue (*post hoc* Sidak's multiple

196   comparisons test, $p=0.0002$), which was also significantly higher than the dart rate to all other

197   cues in the females (*post hoc* Sidak's multiple comparisons test, $p<0.01$). To test if darting

198   behavior in females influenced extinction of freezing, females were categorized into 'darters'

199   and 'non-darters' based on their dart rates exhibited during the last DC session, DC4 (Figure

200   4A). A median split of dart rates during the fear cue in session DC4, created 'darter' (n=10) and

201   'non-darter' (n=10) subgroups. Of the 10 females classified as 'darters', 8 of them also exhibited

202   dart rates in the top 50% during the fear+safety cue (Figure 4A, right). When subdivided,

203   'darters' and 'non-darters' did not show any group differences in extinction of darting (Figure

204   4B). A two-way repeated measures ANOVA of percent time freezing during each fear cue

205   presentation (Figure 4Ci) revealed a group X trial interaction (F(19,342)=1.871, p=0.0154), and

206   main effects of group (F(1, 18)=6.931, p=0.0169) and trial (F(19, 342)=2.621, p=0.0003). When

207   subdivided, 'darters' did show a significant reduction in freezing levels to the fear cue across

208   extinction (*post hoc* Tukey's multiple comparisons test, compared to trial 1, trials 6, 8-10, 14,

209   and 17-20 were significant, p<0.05), while 'non-darters' persisted in their freezing responses

210   throughout all of extinction training (p>0.05 for each trial compared to trial 1). There were also

211   between group differences across extinction training: during trials 15 and 19, 'darters' were

212   significantly lower than 'non-darters' (post hoc Sidak's multiple comparisons test, p<0.05).

213   During the extinction test the next day (Figure 4Cii), between group differences were still

214   detected with 'darters' exhibiting less freezing to the fear cue compared to 'non-darters' (2-way

215   RM ANOVA; group X cue interaction F(3, 54)=4.416, p=0.0075; main effect of cue (F(3,

216   54)=86.36, p<0.0001; *post hoc* Sidak's multiple comparisons test, p=0.0103).

217

218   *Shock reactivity in males versus females*

219   To exclude possible sex differences in pain sensitivity and footshock perception a separate

220   cohort of 8 male and 8 age-matched female rats received 11 unsignaled footshocks of

221   increasing intensities (0.3 mA, 0.35 mA, 0.4 mA, 0.45 mA, 0.5 mA, 0.55 mA, 0.6 mA, 0.7 mA,

222   0.8 mA, 0.9, 1.0 mA) with an inter-stimulus interval of 2 min. Freezing increased as a function of

223   shock intensities (Figure 5A; 2-way RM ANOVA; F(11,121)=25.9, p<0.0001). No main effects of

224   sex (F(1,121)=0.2871, p=0.6027) or sex by shock (F(11,121)=1.413, p=0.1754) were observed.

225   Our experiments utilized a shock intensity of 0.5mA throughout this study. For this particular

226   intensity, we also noted the number of rats that jumped or darted in response to a 0.5mA shock.

227   No sex differences in the number of rats jumping in response to the 0.5mA footshock were

228   observed (χ2: p>0.9). The number of female rats darting after the 0.5mA footshock was higher

229   than males, but not significantly (χ2: p =0.0769), with five of the eight female rats tested

230   exhibiting the behavior. A higher number of females darting in response to the footshock in this

231   test would still not explain the lack of conditioned inhibition of freezing in the females, as

232   freezing levels at 0.5mA was slightly lower than the males (Figure 5A). Our results do not

233   definitively show, but do suggest, that females may respond more preferentially to a footshock

234   with a darting response.

Discussion

In this study, we show females exhibit a significantly different behavioral profile than males in a task that tests for reward, fear and safety cue discrimination, as well as conditioned inhibition and extinction. Female Long Evans rats showed more reward seeking early in training and persistently high freezing levels to the fear cue when in the presence of a safety cue or after fear extinction. When freezing data was re-assessed in the context of darting levels, females categorized as 'darters' did show a significant reduction in freezing levels to the fear cue during extinction, while 'non-darters' did not show any evidence of fear extinction. This data adds to the growing body of evidence of sex differences in fear regulation and highlights the advantages of using more complex learning paradigms with additional behavioral measurements.

Even though studies including female subjects have been proportionally low, several studies have reported clear sex differences in fear regulation. Most of these are consistent with our findings of reduced discrimination between fear and safety signals. For instance, female mice show more generalization of fear to novel and safe contexts compared to males, and with this generalization there is a concurrent increase in basal amygdala activity (Keiser et al. 2017). Male and female rats also respond differently to the controllability of a stressor. Males display reduced fear during escapable stress versus inescapable stress whereas females exhibit no beneficial effects of perceiving a stressor as escapable and controllable (Baratta et al. 2018). The buffering effects seen in these males were linked to prelimbic cortical neurons projecting to the dorsal raphe nucleus, which do not appear to be engaged in females. Females displaying a similar fear response to both inescapable and escapable stress is similar to our findings of females showing equivalent freezing levels to the fear cue in the presence or absence of a safety cue, in that there were no buffering effects seen by the safety cue. It appears that females do not downregulate their fear response in situations cued as safe.

Our data showing an increase in darting behavior in female rats as the number of fear cue-footshock trials increase is consistent with another report using female rats in a fear conditioning and extinction paradigm (Gruene et al. 2015). Like us, Gruene et al (2015) also show darting levels increase as learning about the fear cue advances. Compared to us, Gruene et al (2015) report notably higher darting frequencies, which is most likely due to the differences in shock intensities and number of trials; our study used 4 trials of 0.5mA per day for 4 days compared to their study using 7 trials of 0.7mA on one day. Our study also includes reinforced reward trials within the same sessions as the fear cue-footshock trials, which could alter the contextual

9

269 expectations of the training session and reduce darting levels. Like Gruene et al (2015) we

270 similarly categorized our females into 'darters' and 'non-darters', and both of our studies reveal

271 'darters' exhibiting reduced freezing levels in response to extinction. Interestingly in our study

272 the 'darters' did not significantly reduce darting levels during extinction training but instead they

273 reduced their freezing levels. It would be interesting in future studies to identify what leads a

274 female to become a 'darter' versus 'non-darter'. As darting is a more active response compared

275 to freezing, the circuits engaged during potential threats would likely be different in these two

276 populations.

277

278 Our findings showing a lack of conditioned inhibition of freezing in females appear to be

279 inconsistent with a recent study demonstrating a lack of sex differences in conditioned inhibition

280 of freezing (Foilb et al. 2018). This is likely due to differences in our respective protocols. First,

281 their footshock intensity was 1.2mA, resulting in freezing levels >90% during the fear cue. As

282 footshock intensity and number of trials are consistently inconsistent across studies, it would be

283 interesting to assess if freezing and darting levels in females follow a linear trend with increasing

284 training intensity, or if there is instead a possibly U-shaped relationship. Foilb et al (2018) also

285 used separate presentations of the fear cue and safety cue throughout training and employed

286 the fear+safety cue summation test during recall, whereas we include fear+safety trials as part

287 of the training. In contrast, another study has shown females discriminate equally to males early

288 in training but then generalize their fear response to the safety cue with continued training (Day

289 et al. 2016). While the females in our study clearly showed equivalent freezing levels to both the

290 fear and fear+safety cues at all time points throughout training, they did not increase their

291 freezing levels to the safety cue when presented alone. And, lastly, our paradigm, unlike others,

292 includes reinforced reward trials during the training of fear and safety cues, which would change

293 the context from a 'threat-no threat' situation to a 'threat-no threat-reward' situation, inducing

294 approach behaviors on top of defensive behaviors.

295

296 Altogether, the data paints a consistent picture of females showing heightened fear responses

297 to cues signaling safety, mimicking the clinical picture in women (Gamwell et al. 2014; Lonsdorf

298 et al. 2015). The presentation of a safety signal not only decreases fear, but also stimulates

299 opposing neuronal activity. Viewing synaptic responses in the striatum during safety signal

300 presentation has shown that brain regions dealing with approach and reward become activated

301 (Rogan et al. 2005). These findings have also been translated to using safety signals to

302 overcome anhedonia in rats (Pollak et al. 2008), showing that safety signals may also be

10

303 regulating emotion in addition to conditioned behavior (Foilb & Christianson 2018). In our study,

304 females consistently showed elevated reward seeking behavior during the reward cue

305 compared to males beginning in the second reward pre-training session. This data appears

306 consistent with reward studies showing significant sex differences in response to drugs of

307 abuse, with female rats consistently self-administering drugs more rapidly than males (Becker &

308 Koob 2016). The increased reward seeking in females seen in our study remained until the end

309 of the first DC session at which point they were equivalent to the males. Interestingly, DC1 is the

310 first time the animals are exposed to footshock. Taking into account the lack of conditioned

311 inhibition of freezing in the females, the females may no longer be as motivated to seek rewards

312 in the face of adverse footshocks. This would be consistent with the report that female rats

313 sacrifice their metabolic needs in order to avoid shocks more than males (Pellman et al. 2017).

314

315 Numerous sex differences have been reported in the functioning of the stress neuropeptide,

316 corticotropin-releasing factor (CRF), with differences in receptor expression, distribution,

317 trafficking and signaling (reviewed in (Bangasser & Wiersielis 2018)). The majority of these

318 differences lead to enhanced CRF efficacy in females which may lead to heightened sensitivity

319 to stressors in females. Recently, the gene for CRH receptor 1 (*CRHR1*) has been identified as

320 a possible candidate gene for mood and anxiety disorders. Weber et al. (2016) have shown that

321 carrying the *CRHR1* minor rs17689918 allele increases the risk for panic disorders in women.

322 Patients carrying this risk allele also demonstrate more generalization of fear to a safety cue,

323 increased amygdala activation during the safety cue and decreased frontal cortex activation with

324 discriminative fear conditioning. Thus, aberrant CRF signaling can lead to sustained fear under

325 conditions cued as safe and can be manifested by changes in neural activity in the amygdala

326 and frontal cortex.

327

328 Neural activity in the amygdala and prefrontal cortex has been shown by our lab to also play a

329 critical role in effective discriminative conditioning in male rats. We have previously identified

330 neurons in the basolateral amygdala (BLA) that discriminate among safety, fear and reward

331 cues in male rats (Sangha et al. 2013); our future experiments will test if females show the

332 same discriminative neurons. Using reversible pharmacological inactivations in male rats, we

333 have also demonstrated that the infralimbic prefrontal cortex (IL) is necessary for suppression of

334 conditioned fear during a safety cue and the prelimbic prefrontal cortex (PL) is necessary for

335 fear expression and discriminatory reward seeking (Sangha, Robinson, et al. 2014). These

336 results indicate that activating the IL in the females may improve conditioned inhibition to the

337 combined fear and safety cues. Our results with male rats also show that a D1 receptor agonist

338 infused into the BLA disrupts suppression of conditioned fear (Ng et al. 2018), implicating

339 dopaminergic ventral tegmental area (VTA) neurons projecting to the BLA in safety-fear-reward

340 discrimination. The VTA, and another reward-related area the nucleus accumbens, have been

341 shown to have increased activity in female mice that also show enhanced conditioned place

342 preference for cocaine during oestrus compared to females in dioestrus or males (Calipari et al.

343 2017).

344

345 Our findings are consistent with human studies where females show less discrimination

346 between the fear and safety signals than males (Gamwell et al. 2014; Lonsdorf et al. 2015),

347 which may reflect underlying mechanisms of increased prevalence for anxiety and stress-

348 related disorders in women. For example, a deficiency in effective safety signal processing has

349 been linked to Post-traumatic Stress Disorder (Jovanovic et al. 2009; Jovanovic et al. 2010),

350 panic disorder (Gorka et al. 2014), and anxiety (Lissek et al. 2005), all disorders with a higher

351 incidence in women than men (Mclean et al. 2011). In our paradigm, females show a

352 significantly different behavioral profile than males that is consistent with the clinical picture,

353 thus making it a great tool to test for the neurobiological mechanisms underlying these sex

354 differences.

Methods

Subjects

A total of 24 adult male (215-375g) and 28 adult age-matched female (198-230g) Long Evans rats (Blue Spruce; Envigo, Indianapolis), were single-housed and handled for 1 week prior to testing. All procedures were performed during the light cycle and approved by the Purdue Animal Care and Use Committee. Rats had *ad libitum* access to food and water prior to the start of the experiment. After experiment onset, they were maintained on a food restricted diet (20g per day for males; 16g per day for females) until the last day of the experiment.

Apparatus

The rats were trained in operant conditioning chambers consisting of Plexiglas boxes (32cm length x 25cm width x 30cm height) encased in sound-attenuating chambers (Med Associates, ST Albans, VT). 10% liquid sucrose was delivered through a recessed port 2cm above the floor in the center of one wall. Two lights (28V, 100mA) were located 10.5cm from floor on either side of the port. A light (28V, 100mA) 27cm above the floor on the wall was on throughout the entire session. Auditory cues were delivered via a speaker (ENV-224BM) located 24cm from the floor on the same wall as the port. Footshocks were delivered through a grid floor via a constant current aversive stimulator (ENV-414S). An overhead video camera and side video camera recorded the sessions for subsequent offline video scoring.

Behavioral Procedures

Reward pre-training (5 sessions): An auditory cue is paired with 10% sucrose solution delivery (100µl) and serves as the reward cue (25 trials; ITI, 90-130s).

Habituation (1 session): Rats continue to receive 25 reward cue-sucrose pairings (ITI, 90-130s) in addition to 5 unreinforced presentations ($\varnothing$) each of the future fear and safety cues in order to habituate the rats to their presentation, thereby reducing any baseline freezing to these novel cues.

Discriminative conditioning (4 sessions): Reward cue-sucrose pairings continue (15 trials). Another auditory cue is paired with a mild 0.5mA, 0.5s footshock and serves as the fear cue (4 trials). In separate trials the 20s fear cue is presented at the same time as a 20s safety light cue resulting in no footshock ('fear+safety', 15 trials). Trials in which the safety cue is presented alone without any footshock are also included to assess whether freezing develops to the safety cue as well as providing the animal with additional trials that contains a safety cue-no shock contingency (10 trials). Trials are presented pseudorandomly (ITI, 100-140 s). Cues were

13

389   counterbalanced across subjects with the caveat that the fear+safety compound cue is

390   composed of one auditory cue and one light cue. Eight of the male rats and 12 of the female

391   rats underwent the DC training in which the reward cue was a continuous auditory cue (3 kHz,

392   20s cue; 70dB), the fear cue a pulsing auditory cue (11 kHz, 20s; 70dB) succeeded by a mild

393   footshock (0.5mA, 0.5s), and the safety cue (a light, 28V, 100mA located on both sides of the

394   port) was followed by the absence of a stimulus. The remaining eight male rats and eight female

395   rats underwent training in which the fear and safety cue stimuli were counterbalanced: the light

396   served as the fear cue and the pulsing auditory cue served as the safety cue.

397

398   To exclude possible sex differences in pain sensitivity and foot shock perception a separate

399   batch of male (n=8) and age-matched female (n=8) rats was presented with a series of

400   unsignalled foot shocks of increasing intensities (0.3 mA, 0.35 mA, 0.4 mA, 0.45 mA, 0.5 mA,

401   0.55 mA, 0.6 mA, 0.7 mA, 0.8 mA, 0.9, 1.0 mA) with an inter-stimulus interval of 2 min. The

402   session was flanked with 5 min intervals in which no stimulus occurred.

403

404   *Data analyses*

405   Our experimental groups to directly compare males and females on discrimination behavior

406   consisted of 16-20 rats, which is typical and consistent with our prior published studies using

407   this behavioral task. Cohorts of 4 or 8 female rats were trained alongside cohorts of 4 male rats

408   for a total of 4 replications. Fear behavior was assessed manually offline from videos by

409   measuring freezing, defined as complete immobility with the exception of respiratory

410   movements, which is an innate defensive behavior (Blanchard & Blanchard 1969; Fendt &

411   Fanselow 1999). The total time spent freezing during each 20s cue was quantified and

412   expressed as percentage. Measuring the total time the animal spent inside the reward port and

413   at the entrance of the port with nose positioned at port entrance during each cue assessed

414   reward seeking behavior and was expressed as a percentage. Darting behavior was detected

415   and quantified offline from videos recorded from overhead cameras via a custom MatLab

416   program (Supplementary Files), with movements of a velocity of 23.5cm/s or faster qualifying as

417   a single dart (Gruene et al., 2015); these were also confirmed manually. Darting was expressed

418   as a dart rate, defined as the number of darts per 20s cue. Three individuals performed manual

419   offline behavioral scoring. Pearson's correlations of behavioral values between scorers were

420   greater than $r = 0.80$. Behavioral data were analyzed with one-way or two-way repeated

421   measures ANOVAs, with sex as the independent factor and condition as the repeated factor,

422   followed by *post hoc* Sidak's, Tukey's or Dunnett's multiple comparisons tests with GraphPad

423   Prism 7.

424

425   For shock sensitivity testing, freezing duration in the 2-min intervals between shock

426   presentations was scored manually as an indicator of fear, as well as darting and jumping as an

427   immediate shock response. To compare these simpler reactionary behaviors to differing shock

428   intensities, 8 rats per sex was determined to be sufficient based on similar shock behavior

429   studies. For the freezing durations, a two-way repeated measures ANOVA was carried out via

430   GraphPad Prism 7, with sex as the independent factor and shock intensity as the repeated

431   factor. Darting and jumping were assessed as dichotomous variables with darting/no darting

432   and jumping/no jumping, respectively. For both, a Cochran test was performed.

433

434   **Acknowledgements:**

438

439   **Author contribution:**

440   EMG, MRN, and IM performed experiments, scored behavior and analysed results. KHN

441   created MatLab analysis programs. EMG contributed to writing. SS designed the study,

442   analysed results, and wrote the manuscript.

443

444   **Competing interests:**

445   The authors declare no competing interests.

References

446 Bangasser, D.A. & Wiersielis, K.R., 2018. Sex differences in stress responses : a critical role for
448   corticotropin-releasing factor.

449 Baratta, M. V et al., 2018. Behavioural and neural sequelae of stressor exposure are not
450   modulated by controllability in females. , 47(July 2017), pp.959–967.

451 Becker, J.B. & Koob, G.F., 2016. Sex Differences in Animal Models : Focus on Addiction. ,
452   (April), pp.242–263.

453 Blanchard, R.J. & Blanchard, D.C., 1969. Crouching as an index of fear. *Journal of comparative*
454   *and physiological psychology*, 67(3), pp.370–5.

455 Calipari, E.S. et al., 2017. Dopaminergic dynamics underlying sex-specific cocaine reward.
456   *Nature Communications*, 8.

457 Christianson, J.P. et al., 2012. Inhibition of Fear by Learned Safety Signals: A Mini-Symposium
458   Review. *Journal of Neuroscience*, 32(41), pp.14118–14124.

459 Daviu, N. et al., 2014. Sex differences in the behavioural and hypothalamic-pituitary-adrenal
460   response to contextual fear conditioning in rats. *Hormones and Behavior*, 66(5), pp.713–
461   723. Available at: http://dx.doi.org/10.1016/j.yhbeh.2014.09.015.

462 Day, H.L.L., Reed, M.M. & Stevenson, C.W., 2016. Sex differences in discriminating between
463   cues predicting threat and safety. *Neurobiology of Learning and Memory*, 133, pp.196–203.

464 Fendt, M. & Fanselow, M.S., 1999. The neuroanatomical and neurochemical basis of
465   conditioned fear. *Neuroscience & Biobehavioral Reviews*, 23(5), pp.743–760.

466 Foilb, A.R. et al., 2018. Sex differences in fear discrimination do not manifest as differences in
467   conditioned inhibition. *Learning and Memory*, 25(1), pp.49–53.

468 Foilb, A.R. & Christianson, J.P., 2018. Chapter 11 Brain Mechanisms for Learning and Using
469   Safety Signals. In S. Sangha & D. Foti, eds. *Neurobiology of Abnormal Emotion and*
470   *Motivated Behaviors: Integrating Animal and Human Research*. Cambridge,
471   Massachusetts: Academic Press, pp. 205–224.

472 Gamwell, K. et al., 2014. Fear Conditioned Responses and PTSD Symptoms in Children : Sex
473   Differences in Fear-Related Symptoms.

474 Glover, E.M., Jovanovic, T. & Norrholm, S.D., 2015. Estrogen and extinction of fear memories:
475   Implications for posttraumatic stress disorder treatment. *Biological Psychiatry*, 78(3),
476   pp.178–185. Available at: http://dx.doi.org/10.1016/j.biopsych.2015.02.007.

477 Gorka, S.M. et al., 2014. Aversive responding to safety signals in panic disorder: The
478   moderating role of intolerance of uncertainty. *Journal of Anxiety Disorders*, 28(7), pp.731–
479   736. Available at: http://dx.doi.org/10.1016/j.janxdis.2014.06.004.

480   Grant, B.F. et al., 2016. Epidemiology of DSM-5 drug use disorder results from the national

481   epidemiologic survey on alcohol and related conditions-III. *JAMA Psychiatry*, 73(1), pp.39–

482   47.

483   Gruene, T.M. et al., 2015. Sexually divergent expression of active and passive conditioned fear

484   responses in rats. *eLife*, 4, pp.1–9. Available at:

485   http://elifesciences.org/lookup/doi/10.7554/eLife.11352.

486   Hernandez-Avila, C.A., Rounsaville, B.J. & Kranzler, H.R., 2004. Opioid-, cannabis- and

487   alcohol-dependent women show more rapid progression to substance abuse treatment.

488   *Drug and Alcohol Dependence*, 74(3), pp.265–272.

489   Jovanovic, T. et al., 2010. Impaired fear inhibition is a biomarker of PTSD but not depression.

490   *Depression and anxiety*, 27(3), pp.244–251. Available at:

491   http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2841213&tool=pmcentrez&rend

492   ertype=abstract.

493   Jovanovic, T. et al., 2012. Impaired safety signal learning may be a biomarker of PTSD.

494   *Neuropharmacology*, 62(2), pp.695–704. Available at:

495   http://www.sciencedirect.com/science/article/pii/S0028390811001031.

496   Jovanovic, T. et al., 2009. Posttraumatic stress disorder may be associated with impaired fear

497   inhibition: Relation to symptom severity. *Psychiatry Research*, 167(1–2), pp.151–160.

498   Available at: http://dx.doi.org/10.1016/j.psychres.2007.12.014.

499   Keiser, A.A. et al., 2017. Sex Differences in Context Fear Generalization and Recruitment of

500   Hippocampus and Amygdala during Retrieval. *Neuropsychopharmacology*, 42(2), pp.397–

501   407. Available at: http://dx.doi.org/10.1038/npp.2016.174.

502   Keyes, K.M., Grant, B.F. & Hasin, D.S., 2008. Evidence for a closing gender gap in alcohol use,

503   abuse, and dependence in the United States population. *Drug and Alcohol Dependence*,

504   93(1–2), pp.21–29.

505   Leknes, S. et al., 2011. Relief as a reward: Hedonic and neural responses to safety from pain.

506   *PLoS ONE*, 6(4).

507   Lissek, S. et al., 2005. Classical fear conditioning in the anxiety disorders: A meta-analysis.

508   *Behaviour Research and Therapy*, 43(11), pp.1391–1424.

509   Lonsdorf, T.. B. et al., 2015. Sex differences in CS-discrimination during context-dependent fear

510   learning and its retrieval - the role of biological sex, contraceptives and menstrual cycle

511   phases. *in Press*, 40(6), pp.368–375.

512   Mclean, C.P. et al., 2011. Gender differences in anxiety disorders: Prevalence, course of illness,

513   comorbidity and burden of illness. *Journal of Psychiatric Research*, 45(8), pp.1027–1035.

17

514 Ng, K.H. et al., 2018. Altering D1 receptor activity in the basolateral amygdala impairs fear
515   suppression during a safety cue. *Neurobiology of Learning and Memory*, 147.
516 Pellman, B.A. et al., 2017. Sexually Dimorphic Risk Mitigation Strategies in Rats. *Eneuro*, 4(1),
517   p.ENEURO.0288-16.2017. Available at:
518   http://eneuro.sfn.org/lookup/doi/10.1523/ENEURO.0288-16.2017.
519 Pollak, D.D. et al., 2008. An Animal Model of a Behavioral Intervention for Depression. *Neuron*,
520   60(1), pp.149–161.
521 Rescorla, R.A., 1969. Conditioned inhibition of fear resulting from negative CS-US
522   contingencies. *Journal of comparative and physiological psychology*, 67(4), pp.504–509.
523 Rogan, M.T. et al., 2005. Distinct neural signatures for safety and danger in the amygdala and
524   striatum of the mouse. *Neuron*, 46(2), pp.309–20. Available at:
525   http://www.ncbi.nlm.nih.gov/pubmed/15848808.
526 Sangha, S., Robinson, P.D., et al., 2014. Alterations in reward, fear and safety cue
527   discrimination after inactivation of the rat prelimbic and infralimbic cortices.
528   *Neuropsychopharmacology : official publication of the American College of*
529   *Neuropsychopharmacology*, 39(10), pp.2405–13. Available at:
530   http://www.ncbi.nlm.nih.gov/pubmed/24727732.
531 Sangha, S., Greba, Q., et al., 2014. Heightened fear in response to a safety cue and
532   extinguished fear cue in a rat model of maternal immune activation. *Frontiers in behavioral*
533   *neuroscience*, 8(May), p.168. Available at:
534   http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4019856&tool=pmcentrez&rend
535   ertype=abstract.
536 Sangha, S., Chadick, J.Z. & Janak, P.H., 2013. Safety encoding in the basal amygdala. *The*
537   *Journal of neuroscience : the official journal of the Society for Neuroscience*, 33(9),
538   pp.3744–51. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23447586.
539 Tanimoto, H., Heisenberg, M. & B, G., 2004. Event timing turns punishment to reward. *Nature*,
540   430, p.983. Available at: http://www.nature.com/doifinder/10.1038/nature02816.
541 Walasek, G., Wesierska, M. & Zieliński, K., 1995. Conditioning of fear and conditioning of safety
542   in rats. *Acta neurobiologiae experimentalis*, 55(2), pp.121–132.
543 Weber, H. et al., 2016. Allelic variation in CRHR1 predisposes to panic disorder: Evidence for
544   biased fear processing. *Molecular Psychiatry*, 21(6), pp.813–822.
545

18

Reward seeking			
Session	*Cue x Sex effects*	*Main effect of cue*	*Main effect of sex*
DC1	F(3,102) = 3.472, **p=0.0189**	F(3,102) = 95.16, **p<0.0001**	F(1,34) = 9.827, **p=0.0035**
DC2	F(3,102) = 0.7742, p=0.5110	F(3,102) = 227.9, **p<0.0001**	F(1,34) = 4.69, **p=0.0374**
DC3*	F(3,90) = 0.6512, p=0.5843	F(3,90) = 117, **p<0.0001**	F(1,30)=1.041, p=0.3157
DC4	F(3,102) = 2.255, p=0.0864	F(3,102) = 181.2, **p<0.0001**	F(1,34) = 2.453, p=0.1266
Freezing			
Session	*Cue x Sex effects*	*Main effect of cue*	*Main effect of sex*
DC1	F(3,102) = 2.245, p=0.0876	F(3,102) = 31.82, **p<0.0001**	F(1,34) = 5.045, **p=0.0313**
DC2	F(3,102) = 4.075, **p=0.0089**	F(3,102) = 103.4, **p<0.0001**	F(1,34) = 6.621, **p=0.0146**
DC3*	F(3,90) = 2.9, **p=0.0393**	F(3,90) = 151.3, **p<0.0001**	F(1,30)=9.719, **p=0.0040**
DC4	F(3,102) = 4.889, **p=0.0032**	F(3,102) = 198.9, **p<0.0001**	F(1,34) = 8.294, **p=0.0068**
Darting			
Session	*Cue x Sex effects*	*Main effect of cue*	*Main effect of sex*
DC1	F(3,102) = 1.98, p=0.1216	F(3,102) = 2.388, p=0.0733	F(1,34) = 4.146, **p=0.0496**
DC2	F(3,102) = 1.134, p=0.3390	F(3,102) = 9.377, **p<0.0001**	F(1,34) = 3.667, p=0.0640
DC3*	F(3,90) = 0.6263 p=0.5996	F(3,90) = 19.06, **p<0.0001**	F(1,30)=0.7394, p=0.3959
DC4	F(3,102) = 10.65, **p<0.0001**	F(3,102) = 15.65, **p<0.0001**	F(1,34) = 13.34, **p=0.0009**

*video files for 4 females were corrupted for this session (n=16 females, 16 males)

Table 1. Summary of two-way repeated-measures ANOVA analyses for reward seeking, freezing and darting behaviors during the four discriminative conditioning (DC) sessions.
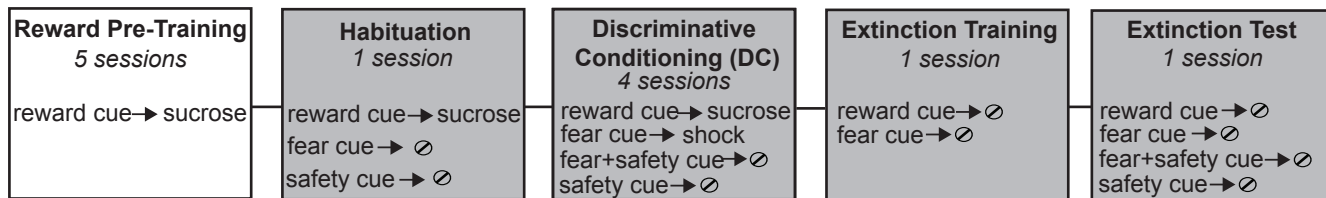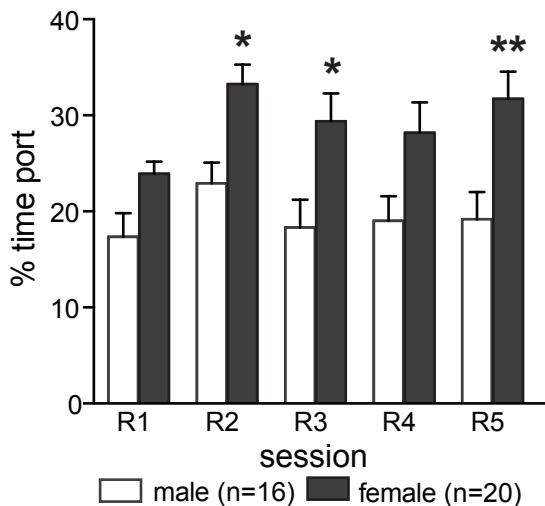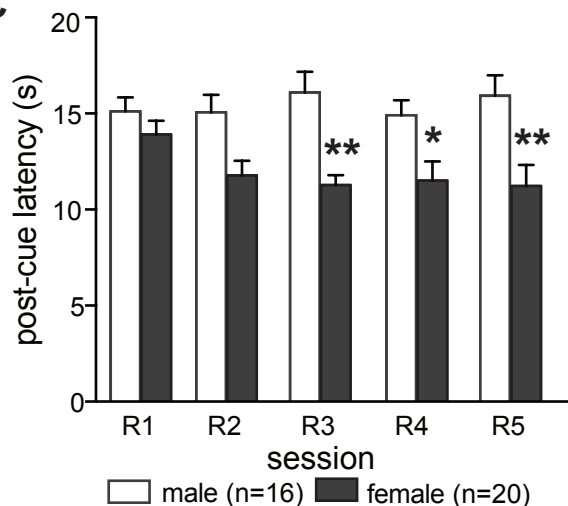
Figure 1. Females show increased reward seeking in response to the reward cue. A) Schematic depicting experimental outline. During reward pre-training, rats (16 males, 20 females) received 25 cue-sucrose pairings across 5 separate sessions. B) Averaged percent time spent in the reward port during the five reward pre-training sessions (R1-5). Females spent significantly more time in the port compared to males during R2, R3 and R5. C) Averaged latency to enter the port after cue onset (in seconds). Females entered the port significantly sooner than males during R3-5. Means +/- SEM. *p<0.05, **p<0.01.
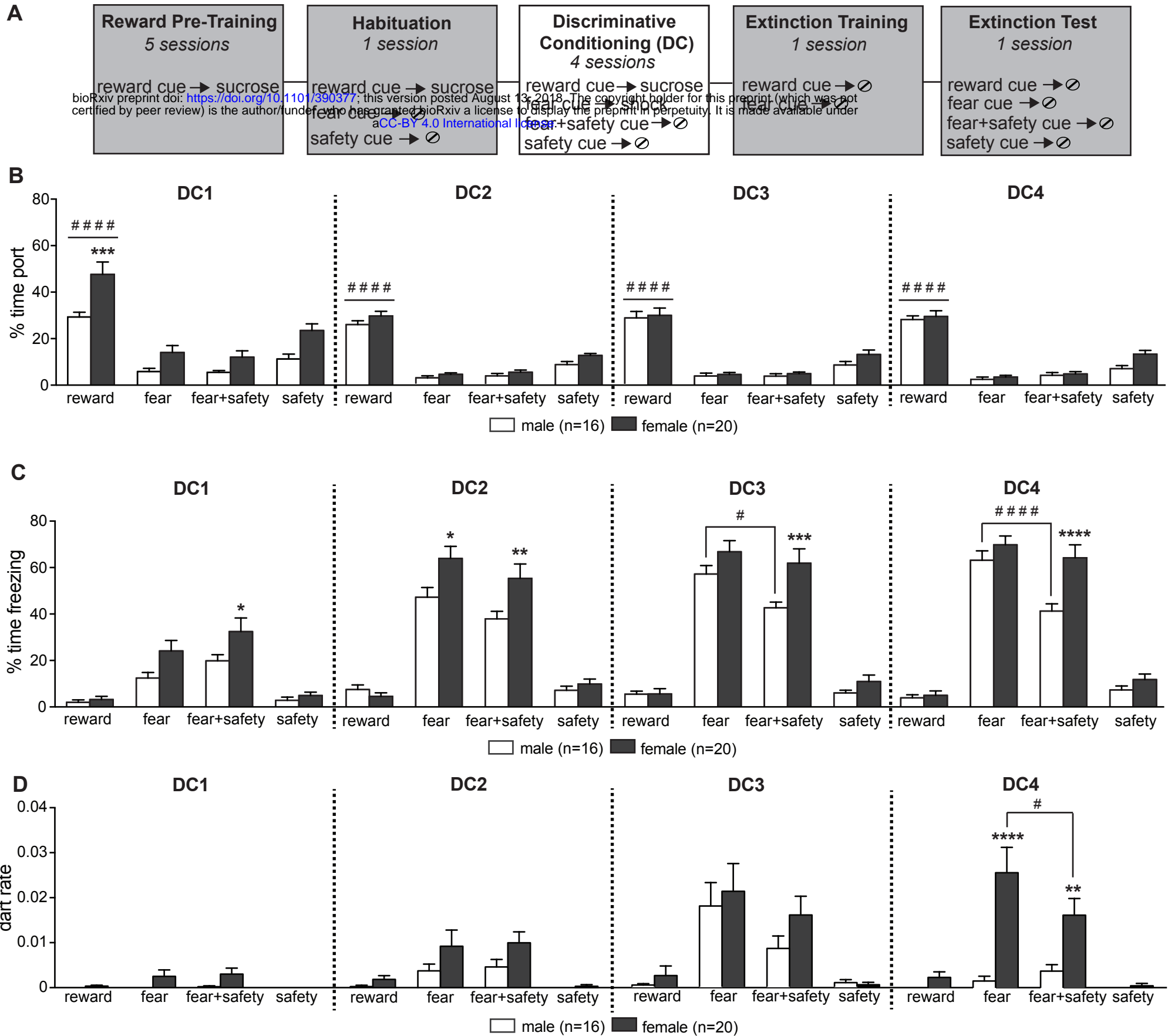
Figure 2. Females do not show inhibition of conditioned freezing in the presence of the safety cue. A) Schematic depicting experimental outline. During the 4 DC sessions, rats (16 males, 20 females) were presented with four types of cued trials: reward cue-sucrose, fear cue-shock, fear+safety cue with no footshock and the safety cue presented alone without footshock. B) Averaged percent time spent in the port during each cue across the 4 DC sessions. Both males and females showed significantly higher reward seeking during the reward cue compared to all other cues during every DC session. During DC1, females showed significantly higher reward seeking to the reward cue compared to males. C) Averaged percent time spent freezing during each cue across the 4 DC sessions. During DC3 and DC4, males (n=16) showed significantly lower freezing to the fear+safety cue (and reward and safety cues) when compared to the fear cue. Females did not show significant inhibition of conditioned freezing to the fear+safety cue compared to the fear cue during any DC session. Females also show significantly higher freezing to the fear+safety cue compared to males during every session. D) Darting behavior during each cue across the 4 DC sessions. During DC4 females showed significantly more darts than males during the fear and fear+safety cues. Also, during DC4 females showed a significant inhibition of darting behavior to the fear+safety cue compared to the fear cue. Means +/- SEM. # p<0.05, ####p<0.0001 within sex, between cue comparison; * p<0.05, **p<0.01, ***p<0.001, ****p<0.0001 within cue, between sex comparison.
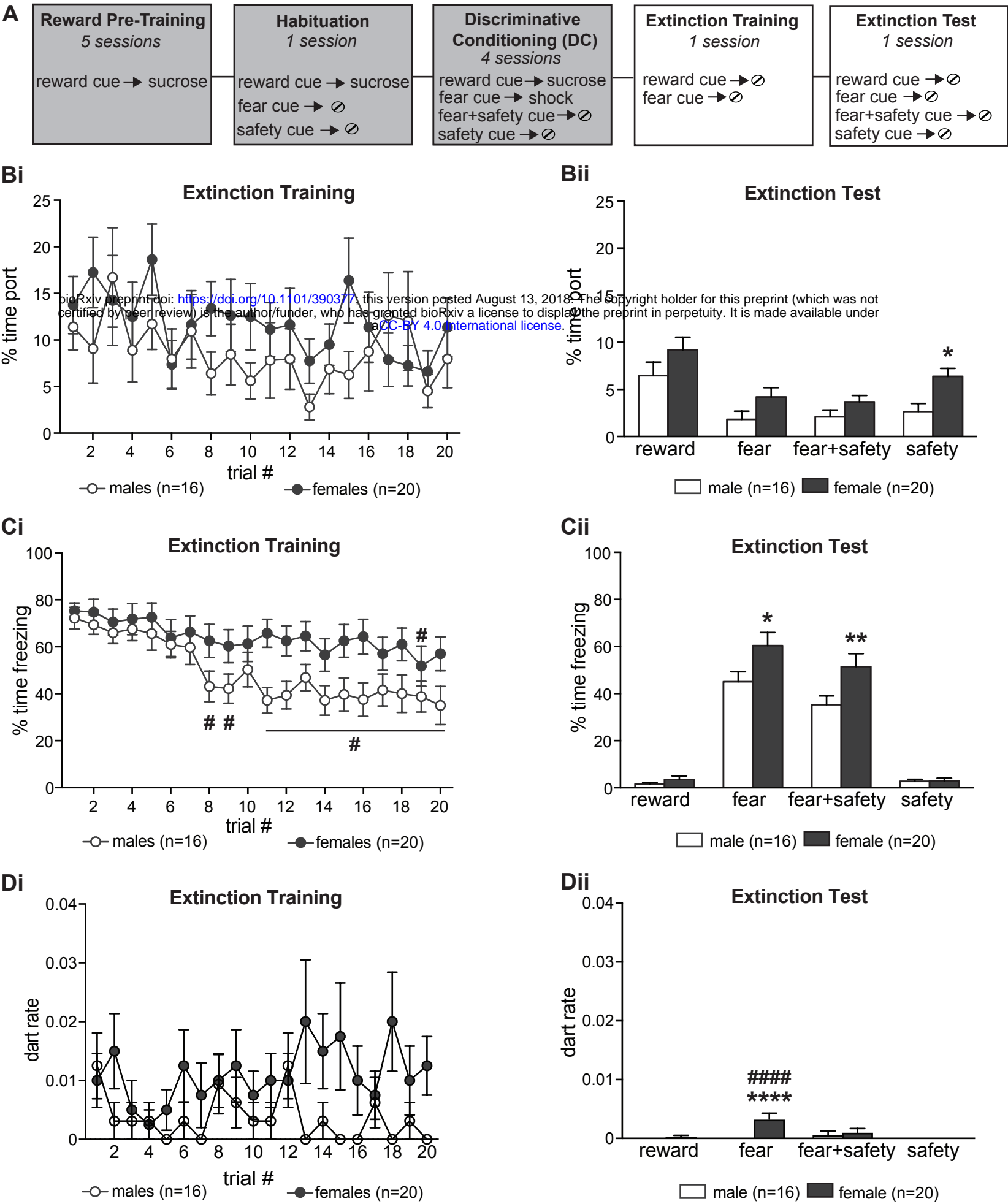
A)

Reward Pre-Training *5 sessions*	Habituation *1 session*	Discriminative Conditioning (DC) *4 sessions*	Extinction Training *1 session*	Extinction Test *1 session*
reward cue → sucrose	reward cue → sucrose fear cue → ⊘ safety cue → ⊘	reward cue → sucrose fear cue → shock fear+safety cue → ⊘ safety cue → ⊘	reward cue → ⊘ fear cue → ⊘	reward cue → ⊘ fear cue → ⊘ fear+safety cue → ⊘ safety cue → ⊘

Figure 3. Females do not show significant extinction of fear. A) Schematic depicting experimental outline. During extinction training both the reward and fear cues are presented in the same session without reinforcement. During the extinction test 1 day later all cues are presented without reinforcement. Bi) Averaged percent time spent in the port during each reward cue presentation during extinction training. No significant differences were found between males and females during extinction training. Bii) Averaged percent time spent in the port during each cue 1 day after extinction training. Females spent significantly more time in the port than males during the safety cue. Ci) Averaged percent time spent freezing during each fear cue presentation during extinction training. Compared to the first trial of extinction, males showed significantly reduced freezing during trials 8, 9 and 11-20. Freezing levels for females did not significantly decrease at any point in extinction training, with the exception of trial 19. #$p<0.05$, compared to trial 1. Cii) Averaged percent time spent freezing during each cue 1 day after extinction training. Males showed evidence of fear cue extinction retention. Females froze significantly more than males during the fear and fear+safety cues. Di) Averaged darting during each fear cue presentation during extinction training. No significant post hoc differences found between males and females during extinction training. Dii) Averaged darting during each cue 1 day after extinction training. Females had a significantly higher dart rate than males during the fear cue. Means +/- SEM. #$p<0.05$, ####$p<0.0001$ within sex, between cue/trial comparisons. *$p<0.05$, **$p<001$, ****$p<0.0001$ within cue, between sex comparisons.

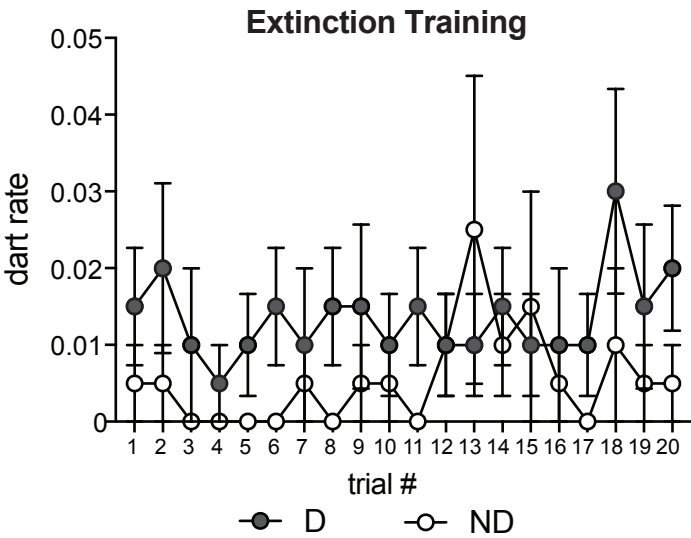**A** Discriminative Conditioning
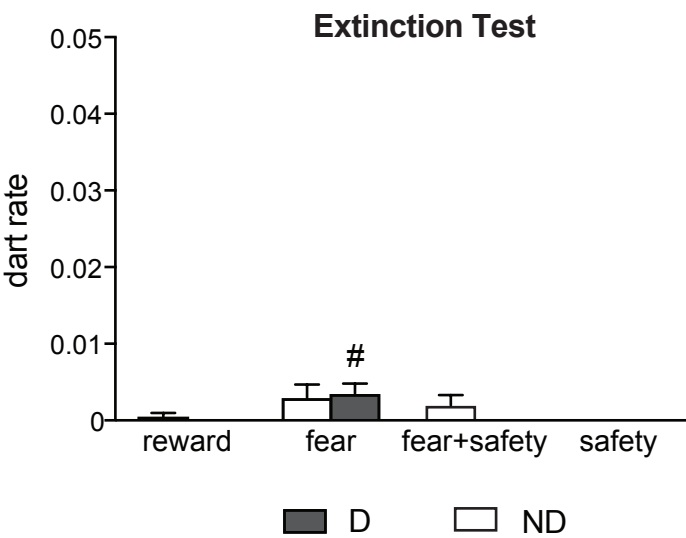*females who dart (D)*
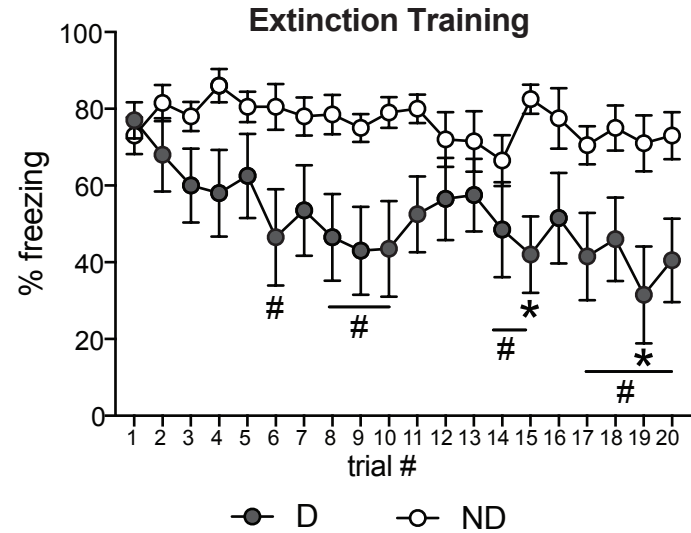*vs*
females who do not (ND)
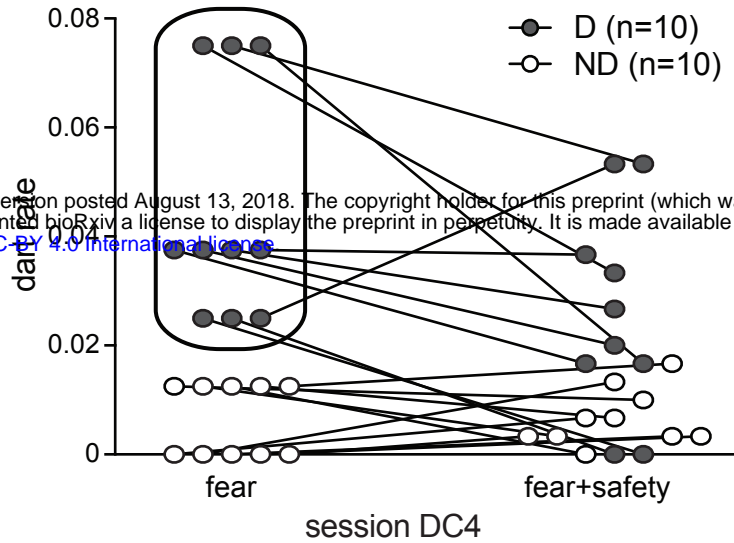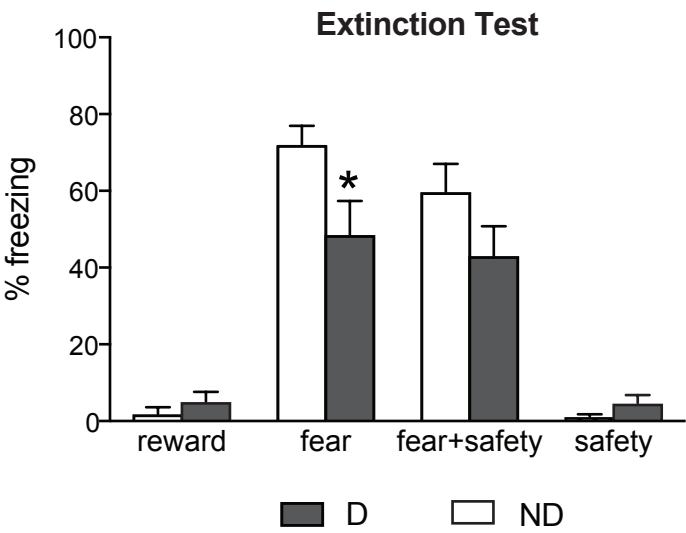
→ ***extinction behavior?***

Figure 4. Extinction differences in female darters versus non-darters. A) Darting behavior during the last discriminative conditioning session was used to split females into 'darters' and 'non-darters' to assess if extinction behavior is influenced. A median split of the number of darts during the fear cue during DC4 was used to categorize female rats as 'darters' and 'non-darters'. For 8 of the 10 females categorized as 'darters' during the fear cue, darting levels were also in the top 50% during the fear+safety cue. Bi, Bii) Averaged dart rates for each fear cue presentation during extinction training. No significant differences in dart rates were seen between groups during extinction training or test. #p<0.05, females had a higher dart rate during the fear cue compared to all other cues. Ci) Averaged percent time spent freezing during each fear cue presentation during extinction training. Compared to the first trial of extinction, females categorized as 'darters' showed a significant reduction in freezing levels during trials 6, 8-10, 14 and 17-20. Freezing levels for females categorized as 'non-darters' did not significantly decrease at any point in extinction training. #p<0.05, compared to trial 1. Cii) Averaged percent time spent freezing during each cue 1 day after extinction training. Darters exhibited less freezing to the fear cue than non-darters. *p<0.05. Means +/- SEM.
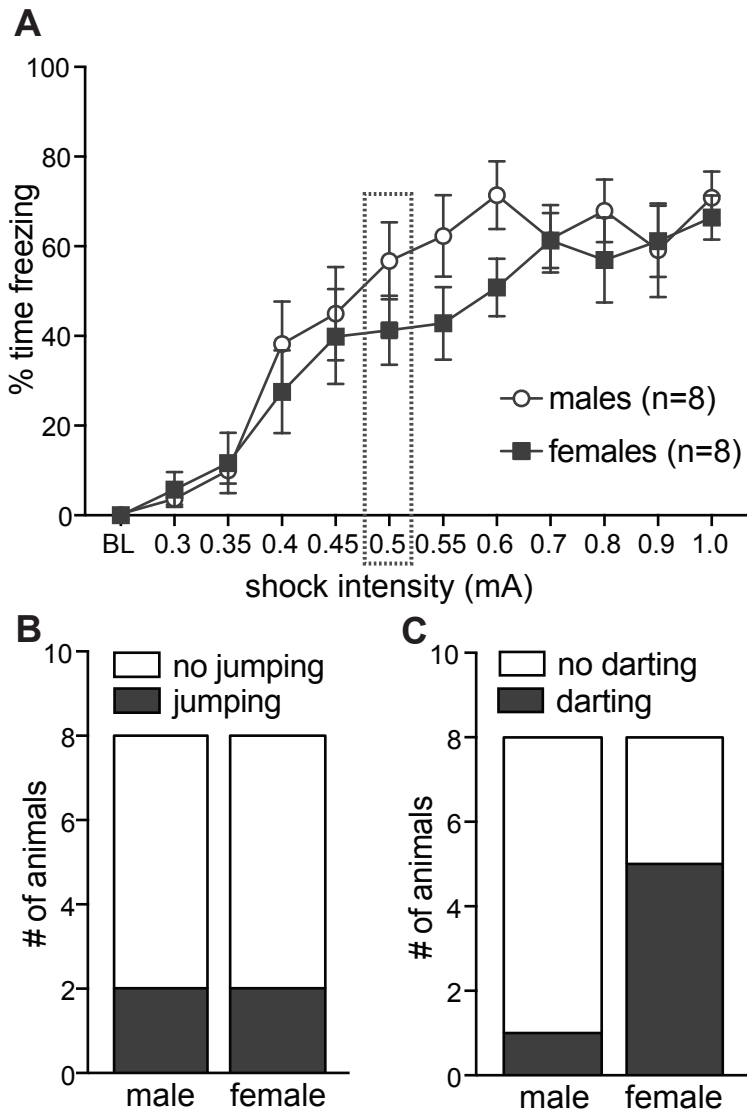
Figure 5. No significant differences in shock reactivity between age-matched male and female rats. A) Male and female rats (n=8 each) were subjected to increasing footshock intensities from 0.3mA to 1.0mA. No significant differences in freezing levels (means +/- SEM) were detected between males and females after each shock presentation. The box around the data at 0.5mA indicates the intensity used for the experiments in this study. There were no significant differences in the number of males or females who jumped (B) or darted (C) in response to the 0.5mA shock.