HHS Public Access

Author manuscript

Behav Brain Res. Author manuscript; available in PMC 2016 April 15.

Published in final edited form as:

Behav Brain Res. 2015 April 15; 283: 121–138. doi:10.1016/j.bbr.2015.01.033.

Testing the role of reward and punishment sensitivity in avoidance behavior: a computational modeling approach

Jony Sheynin[1,2,3,4], **Ahmed A. Moustafa**[1,5], **Kevin D. Beck**[1,2,3], **Richard J. Servatius**[2,3,6], and **Catherine E. Myers**[1,2,3]

Jony Sheynin: jsheynin@med.umich.edu; Ahmed A. Moustafa: A.Moustafa@uws.edu.au; Kevin D. Beck: Kevin.Beck@va.gov; Richard J. Servatius: richard.servatius@va.gov; Catherine E. Myers: Catherine.Myers2@va.gov

[1]Department of Veterans Affairs, New Jersey Health Care System, East Orange, NJ, USA

[2]Joint Biomedical Engineering Program, New Jersey Institute of Technology and Graduate School of Biomedical Sciences, Rutgers, The State University of New Jersey, Newark, NJ, USA

[3]Stress & Motivated Behavior Institute, New Jersey Medical School, Rutgers, The State University of New Jersey, Newark, NJ, USA

[5]Marcs Institute for Brain and Behaviour & School of Social Sciences and Psychology, University of Western Sydney, Sydney, Australia

[6]Department of Veterans Affairs, Veterans Affairs Medical Center, Syracuse, NY, USA

Abstract

Exaggerated avoidance behavior is a predominant symptom in all anxiety disorders and its degree often parallels the development and persistence of these conditions. Both human and non-human animal studies suggest that individual differences as well as various contextual cues may impact avoidance behavior. Specifically, we have recently shown that female sex and inhibited temperament, two anxiety vulnerability factors, are associated with greater duration and rate of the avoidance behavior, as demonstrated on a computer-based task closely related to common rodent avoidance paradigms. We have also demonstrated that avoidance is attenuated by the administration of explicit visual signals during “non-threat” periods (i.e., safety signals). Here, we use a reinforcement-learning network model to investigate the underlying mechanisms of these empirical findings, with a special focus on distinct reward and punishment sensitivities. Model simulations suggest that sex and inhibited temperament are associated with specific aspects of these sensitivities. Specifically, differences in relative sensitivity to reward and punishment might underlie the longer avoidance duration demonstrated by females, whereas higher sensitivity to

Corresponding Author: Jony Sheynin, PhD, Department of Psychiatry, University of Michigan, 4250 Plymouth Rd, Ann Arbor, MI 48109, jsheynin@med.umich.edu, jony.sheynin@gmail.com.
[4]Department of Psychiatry, University of Michigan, Ann Arbor, MI, USA (current affiliation)

Declaration of Interest

The authors affirm that they have no relationships that could constitute potential conflict of interest.

punishment might underlie the higher avoidance rate demonstrated by inhibited individuals. Simulations also suggest that safety signals attenuate avoidance behavior by strengthening the competing approach response. Lastly, several predictions generated by the model suggest that extinction-based cognitive-behavioral therapies might benefit from the use of safety signals, especially if given to individuals with high reward sensitivity and during longer safe periods. Overall, this study is the first to suggest cognitive mechanisms underlying the greater avoidance behavior observed in healthy individuals with different anxiety vulnerabilities.

Keywords

avoidance; computational model; reinforcement learning; anxiety vulnerability; individual differences; safety-signals; cognitive-behavioral therapy

1. Introduction

Avoidance is defined as a behavior that causes the omission of aversive events. Avoidance behavior in response to a cue signaling an upcoming aversive event is usually adaptive and serves to protect one from harm, but exaggerated avoidance behavior is a predominant symptom in all anxiety disorders (American Psychiatric Association, 2000), and its severity parallels the development and persistence of these disorders (North et al. 1999; North et al. 2004; Foa, Stein, and McFarlane 2006; Karamustafalioglu et al. 2006; O'Donnell et al. 2007). To date, the literature on avoidance behavior is based mainly on rodent studies where neutral signals (warning signals; e.g., tones) predict the occurrence of aversive events (e.g., electric shocks), and the animal learns a predetermined response (e.g., lever-press) to overcome these events. Responding during the aversive event results in its termination (escape response; ER), while responding during the warning signal prevents the occurrence of the aversive event (avoidance response; AR).

1.1. Empirical work in human subjects

Some attempts to operationalize human avoidance behavior have used an operant fear-conditioning framework in which the subject makes responses to avoid mild aversive events (“unpleasant but bearable” electric shocks; e.g., Delgado et al. 2009; Lovibond et al. 2008). However, since such stimuli are by definition not highly aversive, the generality of the findings is limited; on the other hand, the use of highly aversive (e.g., painful and distressing) stimuli would have serious ethical and practical constraints. Studies that attempt to address how humans avoid truly painful and/or distressing stimuli have instead tended to rely on self-report questionnaires, which ask subjects to report how often they manifest different types of avoidance behaviors in response to real-world stimuli and events (e.g., Cloninger 1986; Taylor and Sullman 2009). Another line of research employs computer-based tasks to examine avoidance of aversive feedback (e.g., point loss). In these paradigms, the subject controls a spaceship, attempts to gain reward (point gain) by shooting at enemy spaceships, and learns to avoid aversive on-screen events (point loss). These tasks have been successfully shown to assess different aspects of avoidance behavior, such as passive and active avoidance (Arcediano, Ortega, and Matute 1996; and Molet, Leconte, and Rosas 2006, respectively), effects of different reinforcement contingencies and contextual variables

(Raia et al. 2000), as well as discrimination learning and latent inhibition (Byron Nelson and del Carmen Sanjuan 2006).

We have recently extended one of these tasks (by Molet, Leconte, and Rosas 2006) to test the acquisition of escape-avoidance behavior in healthy young adults (Fig. 1; Sheynin et al. 2014a). Briefly, in this task participants controlled a spaceship located at the bottom of the screen and were instructed to maximize their score. Participants could learn that a reward (one point) could be obtained by shooting and destroying an enemy spaceship that was moving on the screen. Every 20 s, a signal (a colored rectangle at the top of the screen) appeared for 5 s. Depending on its color, the signal could be a warning signal (W+) that was followed by an aversive event, or a control signal (W−) that was not associated with any event. The aversive event was a bomb that appeared at the center of the screen for 5 s, during which the participant's spaceship was exploded and a maximum of 30 points could be lost. On warning trials, W+ appeared (warning period), followed by the bomb period, which in turn was followed by a 10-s intertrial interval (ITI) during which no signal/bomb occurred. On control trials, W− appeared (control period), followed by a longer 15-s ITI. Participants could learn to protect themselves from the aversive event by moving their spaceship to a specific "safe area" on the screen ("hiding"). However, while in the safe area, it was impossible to shoot the enemy spaceship and obtain reward. Subjects who entered the safe area during the warning period and remained there throughout the bomb period avoided all point loss on that trial (AR); subjects who entered the safe area after the bomb period began were able to escape that point loss (ER). At the beginning of the experiment, participants were given 1 min of practice time, during which they could shoot the enemy spaceship but no signal or bomb appeared.

Sheynin et al. (2014a) used several variables to describe the escape-avoidance behavior on the computer-based task. First, *hiding duration* indicated the percentage of time spent hiding during the warning period, the control period and the bomb period. Hiding during the bomb period represented an ER and terminated point loss. Hiding during the warning period represented avoidance behavior and could completely prevent any point loss; if the participant emerged from hiding before the end of the bomb period, point loss resumed and response was not recorded as an AR. In addition, Sheynin et al. defined two variables to describe specific aspects of avoidance: *AR rate* – percentage of acquisition trials on which an AR was made and *AR duration* – percentage of the warning period during which the participant's spaceship was hidden, averaged across trials where an AR was made. Longer AR duration indicated that a participant made a response earlier during the warning period and remained hiding longer overall on that trial. In Sheynin et al.'s (2014a) initial study with the spaceship task, the vast majority of the participants learned the ER, while most of them also learned to completely avoid point loss by performing an AR. This pattern is consistent with what is generally reported in the rodent literature on avoidance learning (e.g., Beck et al. 2010).

In addition to providing a framework to operationalize human avoidance behavior, Sheynin et al. (2014a) tested associations of avoidance behavior with individual differences and specifically, those that confer anxiety vulnerability. A large animal literature has demonstrated the effect of strain and sex on active avoidance behavior in rodents.

Specifically, female sex and inhibited temperament (i.e., behavioral inhibition in response to novel or aversive stimuli) have been associated with greater avoidance behavior in rodents (e.g., Beck et al. 2010; Servatius et al. 2008). Since both female sex and inhibited temperament are vulnerability factors for anxiety disorders (Pigott 1999; and Gladstone et al. 2005, respectively), these observations suggested that greater avoidance behavior might mediate vulnerability to anxiety disorders in humans. Indeed, by using the described spaceship avoidance task, Sheynin et al. have found the same facilitated AR pattern in vulnerable young adults. Interestingly, Sheynin et al. (2014a) also reported a double dissociation of sex and temperament. Specifically, although males and females showed similar AR rate, females had longer AR duration, meaning they tended to spent more of the warning period hiding in the safe areas. On the other hand, inhibited participants had higher AR rate than uninhibited participants, with no difference in AR duration. Together, these findings suggested differential vulnerability pathways associated with sex and temperament.

As a follow-up study, Sheynin et al. (2014b) extended the spaceship task to eliminate control trials and to include an extinction phase, where W+ was not followed by an aversive event (bomb and point loss). Importantly, impaired extinction learning characterizes anxiety disorders, as well as post-traumatic stress disorder, and is reflected in patients' tendency to keep emitting ARs, although aversive outcomes no longer occur (Graham and Milad 2011). Results from the acquisition phase on the spaceship task were similar to those of the prior study (Sheynin et al. 2014a), in that females showed longer AR duration than males; females were also slower to extinguish the avoidance behavior than males (shown by longer hiding duration during the warning period on extinction trials), an effect parallel to the delayed avoidance extinction in animal models of anxiety vulnerability (Servatius et al. 2008).

Sheynin et al. (2014b) also used the spaceship task to explore the effect of safety signals (SSs; signals associated with non-threat periods), which were shown to modulate avoidance behavior in rodents (e.g., Beck et al. 2011; Grossen and Bolles 1968; Dillow et al. 1972; for review, see Sheynin et al. 2014b). Participants were divided into two groups given different versions of the spaceship task – "with-SS" and "without-SS". Participants in the "with-SS" group were administered an SS during the ITI on acquisition trials; the SS took the form of two background lights at the two upper corners of the screen (Fig. 1C). Results showed that such signal administration facilitated the extinction of avoidance behavior (shown by decreased hiding during the warning period on extinction trials); this was a main effect that occurred regardless of participant's sex or inhibited temperament.

1.2. Computational modeling approach and the current work

One approach for investigating the mechanisms that might underlie these behavioral outcomes is computational modeling. By developing computational models that simulate the observed behavior, researchers can shed light on previous findings, as well as generate new hypotheses that would drive future empirical work. Traditionally, computational models of avoidance learning have used reinforcement learning algorithms (RL; Johnson et al. 2001; Moutoussis et al. 2008; Maia 2010; Smith et al. 2004; Schmajuk and Zanutto 1997; Myers et al. 2014), in which an "agent" (model) learns by trial and error to maximize reward and minimize punishment (Sutton and Barto 1998). Specifically, if a response to a stimulus

results in a reward (or better-than-expected outcome), the probability of repeating this response in the future is increased. Alternatively, if a response results in a punishment (or worse-than-expected outcome), the probability is decreased. A common computational architecture that describes RL is the actor-critic model (Sutton and Barto 1998). In this model, the "critic" assesses the value of a given state and computes a prediction error (PE), based on the difference between actual and expected outcomes. The "actor" uses this PE to optimize the agent's behavior and maximize the total reward. Normally, RL models include different free parameters, which correspond to specific computational processes in the model and might describe how learning patterns vary across subjects. Parameters include learning rates, exploitation/exploration bias (the tendency to repeat previously-reinforced responses versus explore the effect of new ones; sometimes called "inverse temperature") and sensitivities to the different possible outcomes of the model (e.g., reward and punishment).

Here, we adapt a computational modeling approach that has been previously used to simulate common rodent conditioned avoidance paradigms (e.g., Maia 2010; Myers et al. 2014). Specifically, we use a RL network model with an actor-critic architecture and apply it to the findings from recent studies with the computer-based spaceship task (Sheynin et al. 2014a; Sheynin et al. 2014b). We have organized the results section into three parts (2–4); in each part, we list the key findings in the corresponding empirical work, describe the computational methods that were employed, report the results obtained from the model simulations and discuss the meaning of the results in the context of the existing empirical literature.

Specifically, we first manipulate the different free parameters of the model, with the goal of revealing possible mechanisms that could underlie the associations between female sex, inhibited temperament and facilitated acquisition of avoidance behavior. In light of reports of distinct sensitivities to reward and punishment in females (Li et al. 2007) and inhibited individuals (Torrubia et al. 2001), we hypothesize that different outcome sensitivities may underlie individual differences in avoidance behavior, in which reward and punishment are often competing features (Aupperle et al. 2011). Secondly, we use the model to describe the extinction learning demonstrated on the spaceship task (Sheynin et al. 2014b). We hypothesize that the same mechanisms that would be proposed to underlie females' increased avoidance during the acquisition phase could also underlie their slower extinction learning. Third, we use the model to shed light on the attenuating effect of SS on avoidance behavior (Sheynin et al. 2014b). We hypothesize that since SSs are thought to have rewarding qualities (Christianson et al. 2012), their effect on avoidance behavior could be mediated by an individual's sensitivity to reward. We then employ the model to generate several predictions that could potentially increase the understanding of the involvement of SSs in avoidance behavior, and specifically, as a tool in cognitive-behavioral therapy for anxiety symptoms.

2. Acquisition of avoidance behavior and associations with anxiety vulnerabilities

In the initial study with the spaceship task, Sheynin et al. (2014a) demonstrated that this task could be used to assess the acquisition of escape-avoidance behavior in human subjects. Specifically, although no explicit instructions to use the safe areas were given, healthy young adults successfully learned to discriminate between W− and W+ and to protect themselves from the on-screen aversive event. While the vast majority of the participants learned the ER, some also learned the AR. Importantly, individuals with anxiety vulnerabilities exhibited increased avoidance responding: female sex was associated with longer AR duration and inhibited temperament was associated with higher AR rate.

2.1. Methods

To simulate the spaceship task, we divided each trial of the task (which lasted 20 s) into 60 timesteps, where each timestep represented 0.33 s (Fig. 2). On each timestep, a total of four inputs defined the current state: presence or absence of a warning signal (W+), a control signal (W−), an SS and an aversive event (bomb). To simulate the 1 min of practice time that was given at the beginning of the task, each simulation started with 180 “practice” timesteps during which all inputs were set to zero and only reward was available (i.e., no punishment). Following the methods of Sheynin et al. (2014a), the task then included an acquisition phase that consisted of 24 trials; each trial was preceded by one “pre-trial” timestep where all inputs were set to zero. Then, 15 timesteps with a signal followed; signal type (W+/W−) was determined by the trial type, which followed a pseudorandom but fixed trial order. On warning trials, the signal (W+) was followed by another 15 timesteps with a bomb and 30 ITI timesteps where neither a signal nor a bomb were present (Fig. 2A). On control trials, the signal (W−) was followed by 45 timesteps of ITI (Fig. 2B). On each timestep, external reinforcement could be provided: when the subject’s spaceship was located in the central area during a bomb period, a punishment (R_{punish}) was provided, corresponding to point loss when the subject’s spaceship is bombed. When the subject’s spaceship was located in the central area during any other period, a reward was provided (R_{reward}), corresponding to an opportunity to shoot at the enemy spaceship, which could lead to point gain. When the subject’s spaceship was located in a “safe area”, reinforcement was always zero (no points gained or lost). The subject’s spaceship was placed in the central area on the first timestep of each session (i.e., first timestep of the practice period). R_{punish} and R_{reward} were set as free parameters in the model; since reinforcement values in the current model were arbitrarily set to represent aversive outcomes, R_{punish} and R_{reward} were represented by positive and negative scalars, respectively. The *sensitivity ratio* is defined as the absolute value of the ratio between the values of these outcomes [i.e., abs(R_{punish}/R_{reward})]. On each timestep, the agent could select one of two actions: either to “fight” (remain/move to the central area) and attempt to obtain reward, or “hide” (remain/move to a “safe area”) and avoid possible punishment (red and green arrows in Fig. 2).

In parallel with analyses of the human data from this task, several dependent variables were analyzed. First, on each trial, the percentage of timesteps on which the agent was “hiding” was recorded for the warning, control and bomb periods (*hiding duration*), with hiding

during the warning and bomb period representing avoidance and escape behavior, respectively. Similarly to the human study, we further defined two specific variables: *AR rate* and *AR duration*, to assess the frequency and latency of the avoidance responses. The model simulation recorded an "AR" when the agent entered a safe area during the warning period and remained there through the remaining warning period and the majority (at least 14 timesteps) of the subsequent bomb period. In analogy to the human study, longer AR duration indicated that the agent made a response earlier after onset of W+, and remained in hiding longer overall on that trial.

The critic module—The critic module (see Fig. 3) received four inputs each coding the presence or absence of W+ W−, SS and bomb, and two inputs coding spaceship location (whether it was inside a "safe area" or whether it was in the central area of the screen). At each timestep, the critic computed prediction error (PE), which represented the difference between expected values across adjacent timesteps and was calculated as:

$$PE = R + \gamma * V - V' \quad \text{(Eq. 1)}$$

where R was the reinforcement (R_{punish}, R_{reward} or zero), and γ was the discounting factor that made distant reinforcements count less than more proximate reinforcements; γ was fixed to 0.9 in the current study. V was the predicted future value, calculated as:

$$V = \sum_i v[i] * I_i \quad \text{(Eq. 2)}$$

and V was the value of V from the prior timestep. I_i was the current binary (1/0) value of input i; and $v[i]$ was the strength of connection from input i to V. All $v[i]$ were initialized to zero at the start of the simulation run and updated as:

$$\Delta v[i] = \alpha^{+/-} * PE * I_i \quad \text{(Eq. 3)}$$

where α represented the learning rate (LR), which dictated rate of weight change in the critic; $\alpha^{+/-}$ were the LRs associated with positive/negative PEs (see Frank et al. 2007; 2009). The value of $v[i]$ was clipped at a maximum of R_{punish} and a minimum of zero, to prevent v from growing out of bounds.

The actor module—On each timestep, the actor (see Fig. 3) chose between two possible responses - "fight" or "hide". The probability of selecting a particular response was calculated using a softmax function (Sutton and Barto 1998):

$$P(fight) = f(fight) / [f(fight) + f(hide)] \text{ and } P(hide) = 1 - P(fight) \quad \text{(Eq. 4)}$$

where $f(a) = exp(M_a/T)$, with T being the exploitation/exploration parameter ("inverse temperature") and M_a being the value associated with action a, which was computed as:

$$M_a = \sum_i m[a][i] * I_i \quad \text{(Eq. 5)}$$

As before, I_i was the current value of the input i; $m[a][i]$ was the strength of the connection from input i to action a. To capture the feature that participants were not explicitly informed

about the option of hiding in the safe areas, but had to learn via exploration in the game (Sheynin et al. 2014a), all weights for the "hide" response *m[hide][i]* were initialized to 0.85; since participants were explicitly informed that they should try to gain points by shooting enemy spaceships, the weights for the "fight" response were initialized to a higher value of 0.95. On each timestep, weights for the chosen action *r* were updated based on PE calculated by the critic:

$$\Delta m[r][i]=\varepsilon^{+/-}*(-PE)*I_i \quad \text{(Eq. 6)}$$

where ε represented the LR, which dictated rate of weight change in the actor; $\varepsilon^{+/-}$ were the LRs associated with positive/negative PEs. The values of *m* were restricted to remain positive.

Unless mentioned otherwise, all simulation results represent the average of 100 simulations, which is comparable to number of participants reported in recent studies using this task (Sheynin et al. 2014a; Sheynin et al. 2014b). For simplicity, we first set $(\alpha^{+})=(\alpha^{-})$ and $(\varepsilon^{+})=(\varepsilon^{-})$, as often done with similar models (e.g., Moutoussis et al. 2008; Myers et al. 2014). Then, we show how specific manipulations of these LRs affect model behavior. Except as otherwise noted, parameter values were set to $\alpha^{+/-}=0.0001$, $\varepsilon^{+/-}=0.00021$, $T=0.00133$, R_{punish} = 45 and R_{reward} = −0.9. When values of a specific parameter were manipulated, all other values remained constant. All simulations were run using the Xcode version 4.6.2 programming environment (Apple Inc., Cupertino, CA), using C-source code.

2.2. Results

To test the hypothesis that distinct reward and punishment sensitivities might underlie the associations between anxiety vulnerabilities and AR (Sheynin et al. 2014a), we manipulated the values of these sensitivities in the model and recorded the corresponding change in AR rate and AR duration (Fig. 4). When sensitivity to punishment was increased (R_{punish} range [25,65]) and sensitivity to reward was fixed (R_{reward} = −0.9), both AR rate and AR duration increased – an outcome that did not match any of the empirical findings (Fig. 4A). However, if R_{punish} was increased but the ratio between R_{punish} and R_{reward} was held fixed (sensitivity ratio=50; i.e., R_{reward} was also proportionally increased; Fig. 4B), we observed an increase in AR rate but a minimal change in AR duration – similar to the differences between inhibited and uninhibited individuals observed in Sheynin et al. (2014a). Finally, when the ratio between R_{punish} and R_{reward} was increased (sensitivity ratio range [30,70]) but R_{punish} was fixed (R_{punish} = 45; Fig. 4C), we observed the opposite pattern – an increase in AR duration but a minimal change in AR rate, similar to the differences Sheynin et al. observed between female and male individuals. Based on these observations, we chose parameter values that could represent the different vulnerability groups ["females" versus "males" represented by sensitivity ratio of 65 and 35 (respectively) and R_{punish} of 45 (for both); "inhibited" versus "uninhibited" represented by R_{punish} of 55 and 35 (respectively) and sensitivity ratio of 50 (for both); see vertical lines in Fig. 4B–C].

In the empirical data, inhibited participants had higher AR rate than uninhibited participants (Fig. 5B); however, AR duration did not differ significantly between inhibited and uninhibited participants (Fig 5D). When the value of R_{punish} was set to simulate "inhibited"

and “uninhibited” as described above, the model correctly showed increased AR rate in “inhibited” simulations (Fig. 5A) with little effect on AR duration (Fig. 5C). Further, in the empirical data, males and females did not differ significantly on AR rate (Fig. 5F), but females had significantly longer AR duration than males (Fig. 5H). When the sensitivity ratio was varied to simulate “males” and “females” as described above, the model again approximated the human data: longer AR duration in “female” simulations (Fig. 5G) with smaller effect on AR rate (Fig. 5E).

Lastly, we considered the interaction of sex and temperament. The human participants in Sheynin et al. (2014a) included 22 uninhibited males, 24 uninhibited females, 14 inhibited males and 35 inhibited females. We ran corresponding simulations of “inhibited” and “uninhibited” “males” and “females”, using the parameters illustrated in Fig. 4, and recorded responding for each simulated subject. Data for all 95 simulations were then averaged, to create learning curves similar to those presented for the empirical data (Fig. 6). Just as in the empirical data, the model quickly learned to hide during the bomb period (ER), gradually learned to hide during the warning period (avoidance behavior), with relatively little hiding during control period.

Testing the effect of specific LR manipulations—Distinct punishment/reward sensitivity ratios in males and females (Fig. 4C) could be mediated by striatal dopamine signaling, which is known to play an important role in sensitivity and response to aversive as well as appetitive stimuli (Tomer et al. 2014; van der Schaaf et al. 2014). Specifically, we hypothesized that dopamine D2 receptor binding, which was shown to differ between sexes (Pohjalainen et al. 1998), could be responsible for the different AR duration in males and females. Interestingly, D2 receptor binding has been also associated with the pathway that supports avoidance (“NoGo”) learning, which is triggered by negative PEs in RL models (e.g., Frank et al. 2007; 2009). We thus predicted that manipulating the LR associated with negative PEs in the current model (α^-, ε^-) would affect mainly the simulated AR duration and parallel the reported sex-related difference in avoidance behavior.

Fig. 7 shows AR rate and AR duration as a function of the different LRs. As predicted, increasing ε^- (Fig. 7D) produced a rapid decrease in AR duration (dashed green line) with a much milder decease in AR rate (solid red line). However, increasing α^- produced a similar decrease in both AR duration and AR rate (Fig. 7B), whereas increasing α^+ and ε^+ produced a similar increase in these variables (Fig. 7A and 7C, respectively). Thus, while partially meeting our prediction, these specific manipulations did not adequately address the full range of empirical results.

2.3. Discussion

Here, we demonstrated that a simple actor-critic model can successfully simulate the acquisition of human escape-avoidance behavior, as demonstrated on a computer-based task closely related to common rodent avoidance paradigms. The behavioral paradigm that was simulated here includes two motivational components, namely reward (point gain) and punishment (point loss); the tendency to obtain the rewarding outcome may conflict and compete with the tendency to prevent the punishing outcome (Aupperle et al. 2011). As

hypothesized, model simulations suggest that differential sensitivities to these outcomes result in distinct patterns of avoidance behavior and may shed light on the association between anxiety vulnerabilities and increased AR. Specifically, we show that increased punishment sensitivity might underlie the higher AR rate in inhibited individuals, whereas increased sensitivity ratio (sensitivity to punishment versus reward) might underlie the longer AR duration in females (Sheynin et al. 2014a).

Indeed, the latter finding supports the idea that the balance between punishment and reward sensitivities is as important as these sensitivities themselves (Stein and Paulus 2009). This may be especially true during periods when reward and punishment coincide and tendencies to approach or avoid these outcomes conflict (approach-avoidance conflict; Aupperle et al. 2011). In the spaceship task, the warning period produces competition between the incentive to acquire points by shooting the enemy spaceship (which requires that the participant's spaceship remain in the central area) and the incentive to avoid the upcoming point loss (which requires a hiding response during which the ability to shoot is suspended). Model simulations suggest that the sex difference on AR duration reported by Sheynin et al. (2014a) is the result of different ratios between these competing incentives. Specifically, females' longer AR duration might be the result of higher punishment/reward sensitivity ratio. This is in agreement with prior work showing that female college students reported similar punishment sensitivity but lower reward sensitivity scores than male counterparts (Li et al. 2007; Torrubia et al. 2001). Such a proposition might also provide an explanation for recent reports from both human and non-human animal literature, where females exhibited less approach behavior than males on an approach-avoidance paradigm (Aupperle et al. 2011; and Basso et al. 2011, respectively), as well as on the spaceship task (i.e., less total points gained and less shooting attempts; Sheynin et al. 2014b).

It is also interesting to note that male participants exhibited higher numerical values of AR rate than female participants (Fig. 5F); while this further supports the idea that AR rate and AR duration are distinct and independent types of responding (see opposite patterns in Fig. 5F,H), this trend was supported by neither the model simulation (Fig. 5E), nor by empirical findings in the later spaceship paper (Sheynin et al. 2014b; data not shown) and should be further tested in future work.

While speculation as to the biological causes of the observed sex differences remains beyond the scope of this paper, obvious candidates that might influence reward/punishment sensitivity ratios in males and females might be the different levels of sex hormones (e.g., testosterone; van Honk et al. 2004), as well as other forms of sexual dimorphisms, such as distinct neural activity in brain areas such as amygdala, insula, ventral striatum, prefrontal cortex and/or hippocampus (Aupperle and Paulus 2010; Bach et al. 2014). Interestingly, our results support the idea that distinct dopamine signaling characteristics might also be involved in the different approach/avoidance biases (Pohjalainen et al. 1998; Fig. 7). Specifically, manipulating one LR in the actor (ε^-) provides a close parallel to the sensitivity ratio manipulations (compare Fig. 7D to Fig. 4C), and raises the possibility that D2 receptors in the dorsal striatum (associated with the actor; O'Doherty et al. 2004; Daw 2003) play a predominant role in the sex-related differences in avoidance behavior. Moreover, the association between lower LR and greater AR duration in Fig. 7D is consistent with

females' lower D2 receptor affinity (Pohjalainen et al. 1998), further supporting the importance of the dorsal striatum in the sex differences observed in the current study. However, reports on sex differences in dopaminergic transmission are not consistent and caution should be used when interpreting these results (see also Farde et al. 1995; Parellada et al. 2004; Munro et al. 2006). Future empirical work should specifically study D2 receptor characteristics in the dorsal striatum, test association with avoidance behavior, and use the current computational model to understand the relation to specific cognitive variables such as LRs.

In addition to the association with sex, the personality trait of inhibition affected performance in the spaceship task (Sheynin et al. 2014a). The model suggests that this might be due to higher punishment sensitivity in those with inhibited temperament than in those with uninhibited temperament. This suggestion echoes prior suggestions that approach and avoidance tendencies might be linked to specific personality dimensions (Elliot and Thrash 2002), and is consistent with Gray's biopsychological theory of personality, where a behavioral inhibition system and a behavioral activation system were proposed as the two systems that control behavioral activity (Carver and White 1994). The behavioral inhibition system was thought to become activated by signals of novelty and punishment, and higher activity of this system was shown (by self-report questionnaires) to magnify reactions to negative events (Gable, Reis, and Elliot 2000). Further, the specific relation of the behavioral inhibition system to punishment sensitivity is supported by the positive correlation between self-reported scores that assessed these two constructs in both healthy and clinical populations (Torrubia et al. 2001; and Jappe et al. 2011, respectively). Serotonin, a neuromodulator that was proposed to regulate the behavioral inhibition system (Cloninger 1987), could be involved in punishment sensitivity and should be a target of future work. Overall, the proposition that links inhibited temperament to increased punishment sensitivity provides a simple explanation for the observation that inhibited individuals show facilitation in both operant (Sheynin et al. 2014a; Sheynin et al. 2013) and classical (Myers et al. 2012) conditioning.

Importantly, while the model simulations presented in this part of the work offer plausible mechanisms that are consistent with empirical literature, they address the association between anxiety vulnerability and avoidance behavior only during the acquisition phase. However, greater responding during extinction, when no aversive events occur, is another predominant feature of many psychopathologies (Graham and Milad 2011). In the following section, we adapt the current model to simulate both acquisition and extinction of human avoidance behavior. Due to the failure of specific LR manipulations to parallel differences between inhibited and uninhibited subjects (Fig. 7) and in order to maintain a simple framework, we proceed with the model configuration with fewer free parameters, where $(\alpha^{+})=(\alpha^{-})$ and $(\varepsilon^{+})=(\varepsilon^{-})$. Such configuration is consistent with other related models (e.g., Moutoussis et al. 2008; Myers et al. 2014) and proposes a comprehensive mechanism that is based solely on individual differences in outcome sensitivities. We are specifically interested to test whether the same computational mechanism that was proposed to underlie differences in avoidance acquisition could be applied to address differences in extinction learning.

3. Extending the study to extinction learning

In the second study with the spaceship task, Sheynin et al. (2014b) extended their initial study (Sheynin et al. 2014a) and demonstrated that this task can be successfully used to assess both the acquisition and the extinction of the escape-avoidance behavior in human subjects. Results from the acquisition phase generally replicated the results in Sheynin et al. (2014a): female sex was associated with longer AR duration and inhibited individuals tended to demonstrate higher AR rate, although the latter relationship did not reach statistical significance in Sheynin et al. (2014b). In addition, Sheynin et al. (2014b) also reported that females were slower to extinguish the avoidance behavior than males (shown by longer hiding duration during the warning period on extinction trials).

3.1. Methods

We used the model described earlier to include an extinction phase. Here, we followed the methodology used in Sheynin et al. (2014b). First, all control trials (with W−) were removed from the acquisition phase and only the 12 warning trials (with W+) were left. These were followed by twelve extinction trials, which consisted of 15 timesteps with W+, followed by 45 ITI timesteps (with no signals/bombs; Fig. 8).

3.2. Results

Fig. 9 shows avoidance and escape behavior over the 12 acquisition trials, as well as avoidance behavior during extinction trials (trials 13–23). In the current, as well as in the next part of this study, in all figures that describe responding across both the acquisition and extinction phases (i.e., Fig. 9–13), a grey vertical line represents the end of the acquisition phase. The model successfully accounted for both acquisition and extinction of the avoidance behavior, as reported in Sheynin et al. (2014b). Specifically, simulated females (simulated by increased sensitivity ratio) showed facilitated avoidance acquisition, slower extinction and similar rates of ER, compared to male counterparts (Fig. 9). In addition, in spite of the change in the task design (omission of the control trials), the associations between inhibited temperament, female sex, AR rate and AR duration were replicated (similar to Fig. 5; simulation data not shown).

3.3. Discussion

Impaired extinction learning is a key feature of anxiety disorders, where individuals continue to avoid fear-provoking situations even in the absence of actual threat (Graham and Milad 2011). Interestingly, recent empirical work suggests that, similarly to animal models for anxiety vulnerability (Servatius et al. 2008), healthy humans with anxiety vulnerability due to female sex exhibit slower extinction learning than males (Sheynin et al. 2014b). Model simulations presented here provide a possible interpretation for these empirical findings, and suggest that the same mechanism that successfully accounted for females' facilitated acquisition learning (higher sensitivity ratio, see Fig. 5) can also account for their slower extinction learning (Fig. 9).

These model simulations continue to replicate the finding of longer AR duration in females than males (Sheynin et al. 2014a), as shown in Fig. 5. As in Fig. 5 and in Sheynin et al.

(2014a), the model also continues to predict higher AR rate in inhibited than uninhibited participants. This suggests that, although the association between inhibition and AR rate did not reach significance in Sheynin et al. (2014b), this may be merely due to sampling error, rather than reflecting differences in task design (the omission of control trials); further empirical studies will be needed to clarify this issue.

Further, it is possible that more salient variations in the task design would affect avoidance behavior. For instance, a large literature discusses the involvement of cues that predict the nonoccurrence of an aversive event (i.e., SSs) in aversive learning in rodents, nonhuman primates and humans (for review, see Christianson et al. 2012). Using the current avoidance task, Sheynin et al. (2014b) have recently demonstrated that the administration of an explicit visual SS during the acquisition phase of the task attenuated avoidance behavior. In the next section, we use the computational model to simulate the administration of SSs with the goal of revealing the mechanism that underlies their effect on behavior.

4. Testing the effect of safety signals

In the second study with the spaceship task, Sheynin et al. (2014b) also tested the effect of SSs on avoidance behavior. Results showed that administering SSs during the ITI on acquisition trials facilitated the extinction of the avoidance behavior (shown by decreased hiding during the warning period on extinction trials).

4.1. Methods

We used the model described in part 3.1 to simulate the administration of a signal associated with non-threat periods (i.e., SS). To parallel the procedure used by Sheynin et al. (2014b), we simulated SS administration during the ITI on all acquisition trials, as well as during the initial “practice” period. The SS was simulated by switching the value of the “safety signal” input in the model (input #3; see Fig. 3) to “1” during all corresponding timesteps. We first showed that the model correctly simulated behavioral differences between the two experimental groups (“with-SS” versus “without-SS”). We then investigated how connection strengths in the critic module (dashed blue arrows in Fig. 3) changed as a result of the SS administration, to better understand the computational processes that underlie the simulated behavior. These connection strengths determine the predicted future value V and are associated with each input to the agent (Eq. 2). Based on this expected value, the critic calculates a PE (Eq. 1), which is then used to update the connections in the actor (solid red arrows in Fig. 3). Similarly to the connections in the critic, these connections are associated with the different inputs to the agent and their strengths determine the probability of choosing each action (Eq. 5). Weights were recorded for the 12 acquisition and 12 extinction trials, averaged across all timesteps on each trial. By performing such detailed analyses, we expected to reveal the specific inputs and actions that drove the model behavior.

4.2. Results

In agreement with empirical data, the administration of the SS in the model facilitated the extinction of the avoidance behavior (i.e., decreased hiding during the warning period on extinction trials) without affecting the ER (Fig. 10). This was a main effect that occurred

independent of sex or temperament in both the empirical data and the model (data not shown). There was also an attenuating effect of the SS on avoidance behavior during acquisition; however, this relationship did not reach significance in the empirical data and was weaker than the effect of the SS during extinction in the model.

Interestingly, analyses of the change in the weight of the model connections suggested that the attenuating effect of SS on avoidance behavior is due to the competing approach response (i.e., fighting). Fig. 11A–E show the connection strengths in the critic (top panel; corresponds to the dashed blue arrows in Fig. 3), across acquisition trials 1–12 and extinction trials 13–24. Recall that these weights are used to compute the predicted value (*V*) – i.e., expectation of future reward or punishment, with positive weight values indicating expected punishment and negative values representing expected reward. Considering first the black lines, which represent the "without-SS" condition, Fig. 11A shows that there is a mild increase in weights to the critic from the warning signal during acquisition when the warning signal predicts upcoming threat, followed by a mild decrease in the weight during extinction when warning signal no longer predicts threat (compare to the gradual acquisition and extinction of the avoidance behavior; Fig. 10A,B). Further, Fig. 11C shows that the acquisition is associated with a strong increase in weights to the critic from the aversive event (compare to the dramatic increase in ER; Fig. 10C,D). Note that there is no change in the weight encoding the safety signal (Fig. 11B), since the SS is never presented in the without-SS condition. The final two inputs (Fig. 11D,E) encode location of the participant's spaceship, and both show gradual increases over the acquisition phase, with a mild decrease during extinction.

Turning to the "with-SS" group, there is a similar pattern change in weights to the critic from the warning signal (Fig. 11A), from the aversive event (Fig. 11C) and from the safe area (Fig. 11D) as in the without-SS group. However, since this condition does include presentation of the SS during periods of non-threat, the weights from the SS decrease across acquisition (Fig. 11B). During extinction, the weights from the safety signal do not change further, because the SS is not presented during this phase, and so input #3 is always zero during this phase. More interesting is the difference between with-SS and without-SS conditions in the weight from the central area (Fig. 11E); this weight shows much lower values in the with-SS group, suggesting that the SS decreased the expected punishment associated with the central area.

In addition to changing weights in the critic, weights in the actor (solid red arrows in Fig. 3) are also updated based on prediction error (Eq. 6). There is one set of weights *m["fight"][i]* from each input *i* to the "fight" response, and a second set *m["hide"][i]* from each input *i* to the "hide" response; for simplicity, only the "fight" weights are shown in Fig. 11. Unsurprisingly, the weights to the actor from the various inputs are approximate inverses of those in the critic, since a prediction of upcoming punishment should reduce the likelihood of selecting a "fight" response; thus, for example, after the first few acquisition trials, the aversive event decreases the weight for "fight" (corresponding to a tendency to hide from the bomb; Fig. 11F). Importantly, the SS increases the weight from the central area to the "fight" response, compared to the without-SS condition (Fig. 11J).

The understanding of the mechanisms that result in such weight change is crucial. First, only weights of inputs that are active during a specific period can be updated (see Eq. 3 and 6; when I_i =0, Δv and Δm would also be zero). Second, only weights of chosen actions r are updated (Eq. 6). Third, as different periods in the task often share inputs [e.g., input #6 (being at the central area) can be activated during either the warning period, the bomb period or the ITI], weight alterations during one period could affect model behavior during other periods. In the simulation of the spaceship task, as the subject's spaceship is in the central area during most of the SS appearance (around 92% of the ITI period; data not shown), the corresponding weight from the central area is the main one to be updated (Fig. 11E,J). Such a dramatic increase in the probability to choose the "fight" response when located in the central area during the ITI also affects the warning period, in which subjects spend more than 50% of the time in the central area (see Fig. 10A) – thus, resulting in a decrease of avoidance behavior (decreased hiding during the warning period). Importantly, Fig. 11 presents actor weights that are associated with the "fight" response only; weights associated with the "hide" response were omitted due to a minimal change in their values across the trials (range: 0.8493–0.8559). Such minimal change in probability to hide is due to the fact that no reinforcement is provided when inside a safe area, which in turn results in a small PE (Eq. 1) and accordingly, a small change in probability to hide (Eq. 6). Overall, these analyses suggest that the administration of SS during the ITI reinforced the inputs (SS and the state of being in the central area) and responses ("fight", i.e., stay in the central area) that occurred during that period; since some of these inputs and responses also occurred during other periods (e.g., during the warning period) – behavior during those periods was accordingly affected by the SS.

4.3. Predicting manipulations to the safety signal

We used the model to examine the effect of several manipulations to the SS. First, since common therapeutic approaches for pathological avoidance are based on extinction training (Balooch, Neumann, and Boschen 2012), we tested whether the attenuating effect of SS on avoidance behavior could be obtained if the SS was administered during the extinction phase, instead of the acquisition phase. Fig. 12 shows how administering SS during the ITI on acquisition trials (dotted green line) or on extinction trials (dashed green line) affected the acquisition and extinction of the avoidance behavior (assessed by hiding during the warning period), compared to performance when SS was absent (black solid line). Simulations suggested that administering the SS during the extinction phase did in fact facilitate extinction (Fig. 12).

Second, in light of the important role of the competing appetitive component in modulating the attenuating effect of SS on avoidance behavior (as depicted in Fig. 11), we hypothesized that the sensitivity to reward (R_{reward}) could mediate the SS effect. Simulations showed that administering the SS in a model with different R_{reward} (but fixed R_{punish}) values altered the attenuating effect of SS (Fig. 13): When a low R_{reward} value was used, the effect was minimal, during both the acquisition and extinction phases (Fig. 13A); when medium and high R_{reward} values were used, the effect of the SS was stronger, especially during the extinction phase of the task (Fig. 13B–C).

Third, although a few animal studies have shown that the effect of an SS on avoidance behavior is independent of the SS duration (Galvani and Twitty 1978; Candido, Maldonado, and Vila 1991; Brennan, Beck, and Servatius 2003), the paradigms that were used did not include an explicit appetitive component. Here, we tested the possibility that SS duration could mediate the effect of SS in the spaceship task, which includes an appetitive component (stay in the central area to shoot the enemy spaceship for point gain) that competes with the aversive component (hide and avoid point loss). Fig. 14 shows how administering the SS during the ITI on acquisition trials, in models with different ITI durations (10, 30 and 50 timesteps) modulated the attenuating effect of the SS on avoidance behavior (Fig. 14): Similarly to the effect of different R_{reward} values (Fig. 13), the effect was minimal when short ITI was used (Fig. 14A) and was stronger when medium and long durations were used, especially during the extinction phase of the task (Fig. 14B–C).

4.4. Discussion

Periods free from aversive events (i.e., safety periods) are thought to represent an appetitive component that is capable of modulating avoidance behavior in rodents (Denny and Weisman 1964; Berger and Brush 1975). Moreover, it has been argued that signals that are associated with these periods (SSs) may provide positive reinforcement and may become inhibitors of fear (Christianson et al. 2012). Although a rich rodent literature exists on this topic, the lack of a standardized methodology together with inconsistent results in the rodent literature (for review, see Sheynin et al. 2014b) limit interpretation and translation of work to a human population. In addition, in spite of the importance of extinction learning as therapy for anxiety symptomatology, reports on the role of SSs in extinction are few and inconsistent. To bridge the gap between human and non-human animal literature, we have recently used the spaceship task to test the role of SS in human avoidance behavior (Sheynin et al. 2014b). We found that the administration of a visual SS during ITI (and “practice” period) of the acquisition phase attenuated the demonstrated avoidance behavior. Consistent with these empirical findings, model simulations presented in the current work similarly demonstrate the decreased avoidance in response to SS administration.

While the model successfully replicates prior empirical findings on avoidance behavior, its core value lies in its ability to reveal mechanisms that could underlie the observed outcomes. Specifically, the ITI and the “practice” time are periods when reward is available and no punishment can occur. By administering an SS during these “safe” periods, the critic increased the values associated with the inputs that were active during these periods, corresponding to an expectation of reward. As the critic learned this prediction, weights were adjusted in the actor to favor the actions that were chosen (during these periods). Specifically, the actor increased the probability of “fighting” (remaining in the central area) when the SS was present. It is important to note that since different periods in the task often share inputs [e.g., input #6 (location at the central area) can be activated during either the warning period, the bomb period or the ITI], such weight alteration also affected model behavior during the warning period – the probability of remaining in the central area (“fighting”) was increased, and accordingly, avoidance behavior was decreased. Such a link between SSs and the appetitive component of an avoidance paradigm is consistent with earlier rodent reports, where safe states (e.g., “non-shock areas”) induced “relaxation”,

reduced avoidance, and were generalized to other states with similar cues (Denny and Weisman 1964).

Lastly, we used the model to generate several novel predictions, which would motivate new investigations to better exploit SSs as a vital component in anxiety therapy. First, while the effect of an SS administered during the acquisition phase is consistent with a large rodent literature and is important for promoting basic understanding of avoidance behavior, obtaining control over individual stimuli during real-life situations is not always possible. Thus, administering or removing potential SSs during acquisition of avoidance may not be practical. Cognitive-behavioral therapies, however, often rely on extinction learning, where individuals are exposed to the feared stimulus or outcome in the absence of actual danger (Balooch, Neumann, and Boschen 2012). The model predicts that, similar to the effects of SS administration during the acquisition phase, an SS administered during the extinction phase would similarly lead to facilitation of extinction learning. Second, the model predicts that when reward sensitivity is high, the attenuating effect of the SS is also increased. Interestingly, the model also suggests that when reward sensitivity is low, the SS might have no effect – a finding that might explain some of the inconsistency in the rodent literature, where effect of SSs on avoidance extinction is not always observed (Dillow et al. 1972; Candido, Maldonado, and Vila 1991; Fernando et al. 2014). Third, the model predicts that using longer safety (ITI) periods could help magnify the attenuating effect of SSs on avoidance behavior, especially in paradigms that include an explicit appetitive component.

5. Overall summary and conclusions

In this work we first demonstrated that increased sensitivity to an aversive outcome could account for the higher AR rate demonstrated by inhibited individuals. This idea is consistent with Gray's original definition of a behavioral inhibition system (Carver and White 1994) and provides a simple explanation for the association between inhibited temperament and facilitated avoidance learning on computer-based tasks (Sheynin et al. 2013; Sheynin et al. 2014a). We then used the model to show that higher ratio of punishment/reward sensitivity could underlie the longer AR duration, as well as the slower extinction learning, demonstrated by females (Sheynin et al. 2014a; Sheynin et al. 2014b). Imbalance between the approach and avoidance systems has previously been argued to characterize anxiety disorders (Stein and Paulus 2009), and might partially explain why females are more vulnerable than males to anxiety disorders. Specific manipulations of the LRs also suggest that these sex differences might be mediated by dopaminergic signaling in the dorsal striatum – an idea that is consistent with previous literature and should drive future empirical studies. The model further suggested an interpretation for the finding that SSs are capable of attenuating avoidance behavior (Sheynin et al. 2014b), by increasing the probability of choosing a competing approach response.

The current work has important implications for future behavioral studies. First, the model made several novel predictions that could be tested empirically. We predicted that SSs administered during extinction learning would also retard avoidance behavior, that SSs might have a larger attenuating effect in individuals with high reward sensitivity, and that the effect of SSs on avoidance could be increased by using longer safety periods. The

spaceship task and common rodent avoidance paradigms could be employed in future studies to test these predictions in both human and non-human subjects, respectively.

The model also predicted that females may have different relative sensitivity to reward versus punishment than males, while those with inhibited temperament should have greater sensitivity to punishment than those with uninhibited temperament. This might be explored in humans by self-report questionnaires that would specifically assess participants' approach-avoidance biases. Such questionnaires might include the Sensitivity to Punishment and Sensitivity to Reward Questionnaire (Torrubia et al. 2001), the Behavioral Inhibition/ Activation Scale (Carver and White 1994), or specific ratings of the motivation to approach reward and avoid punishment during the task (Aupperle et al. 2011).

Importantly, the model described here specifically addresses the approach-avoidance conflict introduced by the spaceship task. As such, it uses a simplified environment where two competing responses are available: an approach response, which provides an opportunity to obtain reward ("fight") and an avoidance response, which prevents possible punishment ("hide"). Future empirical and computational work could further examine performance on the task by assessing the degree that such approach behavior translates to an actual reward, i.e., whether the subject chooses to "shoot" while being in the central area, and whether such shooting is accurate and hits the enemy spaceship.

It should also be noted that the current work used a parameter-tuning approach to test specific hypotheses regarding the role of distinct reward and punishment sensitivities in avoidance behavior. While a few models of avoidance behavior have been previously reported (Johnson et al. 2001; Moutoussis et al. 2008; Maia 2010; Smith et al. 2004; Schmajuk and Zanutto 1997), the current model together with another parameter-tuning model of rodent avoidance behavior (Myers et al. 2014) represent the first attempts to use computational techniques to understand individual differences in this behavior. The parameter-tuning approach is often used to test the effect of specific parameters on the model behavior (e.g., Tables 1–2 in O'Reilly and Frank 2006), to simulate group differences (e.g., in Parkinson's disease; Moustafa and Gluck 2011) and is especially advantageous when addressing *a priori* hypotheses concerning possible abnormal processes in the studied subjects (for review, see Fig. 1 in Maia and Frank 2011). Moreover, we predicted that the proposed mechanisms would successfully describe group differences across several experiments; utilizing a model with a fixed set of tuned parameters provides a simplified and unified framework that can potentially explain such a complex empirical dataset. Future computational efforts could further extend the current findings and use a fitting approach to simulate data at the individual level and test whether other specific parametric configurations could offer an alternative explanation (e.g., see Frank et al. 2007; 2009); a synergistic use of different approaches or models at different levels of abstraction should also be considered (Maia and Frank 2011). Overall, it should be clear that rather than providing the best explanation, this work merely proposes one simple and comprehensive mechanism that should be tested by future empirical analyses.

Lastly, future work should build on the empirical and computational grounds presented here to address other important features of avoidance learning, e.g., the role of response-stimulus

contingency in the effect of SSs on avoidance acquisition: While prior rodent studies demonstrated that the presence of an SS facilitated acquisition of AR, Sheynin et al.'s (2014b) study showed the opposite pattern; Sheynin et al. raised the possibility that such discrepancy is due to the contingency of the SS on the rat's ER/AR, whereas in the spaceship task the SS always appeared at the end of the bomb period, regardless of the subject's behavior. Another phenomenon that could be investigated is "warm-up" (decreased AR at the beginning of a new session); lack of warm-up was shown to characterize vulnerable subjects in both empirical rodent work and computational modeling (Servatius et al. 2008; and Myers et al. 2014, respectively). Further, the spaceship task and the computational model could be used to study active versus passive avoidance mechanisms, and specifically, testing the idea that female sex and inhibited temperament might be associated with different types of avoidance (Sheynin et al. 2014b). Lastly, RL models have been linked to brain substrates; specifically, fMRI and electrophysiological studies provided evidence for the idea that the critic (value prediction) is related to the ventral striatum and the actor (action selection) is related to the dorsal striatum in both humans and rodents (O'Doherty et al. 2004; and Daw 2003, respectively). Possible future directions could include functional neuroimaging (e.g., fMRI) studies of humans performing the spaceship task, to reveal the brain areas that are involved in different patterns of the avoidance behavior (e.g., higher AR rate or longer AR duration) in individuals with different anxiety vulnerabilities.

In sum, the current work presents for the first time the use of computational modeling techniques to study the processes that might result in greater avoidance in vulnerable individuals, which in turn might subsequently increase the development of pathological avoidance behaviors. Importantly, by simulating human behavior on a task that was developed to parallel common animal paradigms, this work helps to bridge the gap between human and non-human avoidance literature. The model simulations propose a simple and unified explanation for a series of empirical reports obtained from healthy college students with anxiety vulnerabilities; this explanation is based solely on individual differences in sensitivity to rewarding and punishing outcomes, and as such, may generalize from experimental to real-world settings. Finally, we used the model to generate several predictions regarding how SSs could be used to promote or retard avoidance; if these predictions are upheld, this would provide further support for the computational model and could also provide valuable insight to clinicians seeking to optimize extinction-based therapies for individuals with pathological avoidance.

Acknowledgments

This work was supported by Award Number I01CX000771 from the Clinical Science Research and Development Service of the VA Office of Research and Development, by the NSF/NIH Collaborative Research in Computational Neuroscience (CRCNS) Program, by NIAAA (R01 AA018737), and by additional support from the SMBI. The views in this paper are those of the authors and do not represent the official views of the Department of Veterans Affairs or the U.S. government.

References

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR. 4th Edn. Washington, DC: APA; 2000.

Arcediano F, Ortega N, Matute H. A behavioural preparation for the study of human Pavlovian conditioning. Q J Exp Psychol B. 1996; 49:270–283. [PubMed: 8828400]

Aupperle RL, Paulus MP. Neural systems underlying approach and avoidance in anxiety disorders. Dialogues Clin Neurosci. 2010; 12:517–531. [PubMed: 21319496]

Aupperle RL, Sullivan S, Melrose AJ, Paulus MP, Stein MB. A reverse translational approach to quantify approach-avoidance conflict in humans. Behav Brain Res. 2011; 225:455–463. [PubMed: 21843556]

Bach DR, Guitart-Masip M, Packard PA, Miro J, Falip M, Fuentemilla L, Dolan RJ. Human hippocampus arbitrates approach-avoidance conflict. Curr Biol. 2014; 24:541–547. [PubMed: 24560572]

Balooch SB, Neumann DL, Boschen MJ. Extinction treatment in multiple contexts attenuates ABC renewal in humans. Behav Res Ther. 2012; 50:604–609. [PubMed: 22835841]

Basso AM, Gallagher KB, Mikusa JP, Rueter LE. Vogel conflict test: sex differences and pharmacological validation of the model. Behav Brain Res. 2011; 218:174–183. [PubMed: 21115068]

Beck KD, Jiao X, Pang KC, Servatius RJ. Vulnerability factors in anxiety determined through differences in active-avoidance behavior. Prog Neuropsychopharmacol Biol Psychiatry. 2010; 34:852–860. [PubMed: 20382195]

Beck KD, Jiao X, Ricart TM, Myers CE, Minor TR, Pang KC, Servatius RJ. Vulnerability factors in anxiety: Strain and sex differences in the use of signals associated with non-threat during the acquisition and extinction of active-avoidance behavior. Prog Neuropsychopharmacol Biol Psychiatry. 2011; 35:1659–1670. [PubMed: 21601608]

Berger DF, Brush FR. Rapid acquisition of discrete-trial lever-press avoidance: effects of signal-shock interval. J. Exp. Anal. Behav. 1975; 24:227–239. [PubMed: 16811875]

Brennan FX, Beck KD, Servatius RJ. Leverpress escape/avoidance conditioning in rats: safety signal length and avoidance performance. Integr Physiol Behav Sci. 2003; 38:36–44. [PubMed: 12814195]

Byron Nelson J, del Carmen Sanjuan M. A context-specific latent inhibition effect in a human conditioned suppression task. Q J Exp Psychol (Hove). 2006; 59:1003–1020. [PubMed: 16885140]

Candido A, Maldonado A, Vila J. Effects of duration of feedback on signaled avoidance. Anim Learn Behav. 1991; 19:81–887.

Carver CS, White TL. Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. J Pers Soc Psychol. 1994; 67:319–333.

Christianson JP, Fernando AB, Kazama AM, Jovanovic T, Ostroff LE, Sangha S. Inhibition of fear by learned safety signals: a mini-symposium review. J Neurosci. 2012; 32:14118–14124. [PubMed: 23055481]

Cloninger CR. A unified biosocial theory of personality and its role in the development of anxiety states. Psychiatric Dev. 1986; 4:167–226.

Cloninger CR. A systematic method for clinical description and classification of personality variants. A proposal. Arch Gen Psychiatry. 1987; 44:573–588. [PubMed: 3579504]

Daw, ND. Reinforcement learning models of the dopamine system and their behavioral implications. Unpublished doctoral dissertation. Pittsburgh: Carnegie Mellon University; 2003.

Delgado MR, Jou RL, Ledoux JE, Phelps EA. Avoiding negative outcomes: tracking the mechanisms of avoidance learning in humans during fear conditioning. Front Behav Neurosci. 2009; 3:33. [PubMed: 19847311]

Denny MR, Weisman RG. Avoidance Behavior as a Function of Length of Nonshock Confinement. J. Comp. Physiol. Psychol. 1964; 58:252–257. [PubMed: 14215399]

Dillow PV, Myerson J, Slaughter L, Hurwitz HMB. Safety signals and the acquisition and extinction of lever-press discriminated avoidance in rats. British J Psychol. 1972; 63:583–591.

Elliot AJ, Thrash TM. Approach-avoidance motivation in personality: approach and avoidance temperaments and goals. J Pers Soc Psychol. 2002; 82:804–818. [PubMed: 12003479]

Farde L, Hall H, Pauli S, Halldin C. Variability in D2-dopamine receptor density and affinity: a PET study with [11C]raclopride in man. Synapse. 1995; 20:200–208. [PubMed: 7570351]

Fernando AB, Urcelay GP, Mar AC, Dickinson TA, Robbins TW. 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. Neuropsychopharmacol. 2014; 39:1420–1430.

Foa EB, Stein DJ, McFarlane AC. Symptomatology and psychopathology of mental health problems after disaster. J Clin Psychiatry. 2006; 67(Suppl 2):15–25. [PubMed: 16602811]

Frank MJ, Doll BB, Oas-Terpstra J, Moreno F. Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. Nat Neurosci. 2009; 12:1062–1068. [PubMed: 19620978]

Frank MJ, Moustafa AA, Haughey HM, Curran T, Hutchison KE. Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. Proc Natl Acad Sci U S A. 2007; 104:16311–16316. [PubMed: 17913879]

Gable SL, Reis HT, Elliot AJ. Behavioral activation and inhibition in everyday life. J Pers Soc Psychol. 2000; 78:1135–1149. [PubMed: 10870914]

Galvani PF, Twitty MT. Effects of intertrial interval and exteroceptive feedback duration on discriminative avoidance acquisition in the gerbil. Anim Learn Behav. 1978; 6:166–173.

Gladstone GL, Parker GB, Mitchell PB, Wilhelm KA, Malhi GS. Relationship between self-reported childhood behavioral inhibition and lifetime anxiety disorders in a clinical sample. Depress Anxiety. 2005; 22:103–113. [PubMed: 16149043]

Graham BM, Milad MR. The study of fear extinction: implications for anxiety disorders. Am J Psychiatry. 2011; 168:1255–1265. [PubMed: 21865528]

Grossen NE, Bolles RC. Effects of a classical conditioned ‘fear signal' and ‘safety signal' on nondiscriminated avoidance behavior. Psychon Sci. 1968; 11:321–322.

Jappe LM, Frank GK, Shott ME, Rollin MD, Pryor T, Hagman JO, Yang TT, Davis E. Heightened sensitivity to reward and punishment in anorexia nervosa. Int J Eat Disord. 2011; 44:317–324. [PubMed: 21472750]

Johnson JD, Li W, Li J, Klopf AH. A computational model of learned avoidance behavior in a one-way avoidance experiment. Adap Behav. 2001; 9:91–104.

Karamustafalioglu OK, Zohar J, Guveli M, Gal G, Bakim B, Fostick L, Karamustafalioglu N, Sasson Y. Natural course of posttraumatic stress disorder: a 20-month prospective study of Turkish earthquake survivors. J Clin Psychiatry. 2006; 67:882–889. [PubMed: 16848647]

Li CR, Huang CY, Lin W, Sun CWV. Gender differences in punishment and reward sensitivity in a sample of Taiwanese college students. Pers Individ Dif. 2007; 43:475–483.

Lovibond PF, Saunders JC, Weidemann G, Mitchell CJ. Evidence for expectancy as a mediator of avoidance and anxiety in a laboratory model of human avoidance learning. Q J Exp Psychol (Hove). 2008; 61:1199–1216. [PubMed: 18938780]

Maia TV. Two-factor theory, the actor-critic model, and conditioned avoidance. Learn Behav. 2010; 38:50–67. [PubMed: 20065349]

Maia TV, Frank MJ. From reinforcement learning models to psychiatric and neurological disorders. Nat Neurosci. 2011; 14:154–162. [PubMed: 21270784]

Molet M, Leconte C, Rosas JM. Acquisition, extinction and temporal discrimination in human conditioned avoidance. Behav Processes. 2006; 73:199–208. [PubMed: 16806735]

Moustafa AA, Gluck MA. A neurocomputational model of dopamine and prefrontal-striatal interactions during multicue category learning by Parkinson patients. J Cogn Neurosci. 2011; 23:151–167. [PubMed: 20044893]

Moutoussis M, Bentall RP, Williams J, Dayan P. A temporal difference account of avoidance learning. Network. 2008; 19:137–160. [PubMed: 18569725]

Munro CA, McCaul ME, Wong DF, Oswald LM, Zhou Y, Brasic J, Kuwabara H, Kumar A, Alexander M, Ye W, Wand GS. Sex differences in striatal dopamine release in healthy adults. Biol Psychiatry. 2006; 59:966–974. [PubMed: 16616726]

Myers CE, Smith IM, Servatius RJ, Beck KD. Absence of “warm-up” during active avoidance learning in a rat model of anxiety vulnerability: insights from computational modeling. Front Behav Neurosci. 2014; 18(8):283. [PubMed: 25183956]

Myers CE, Vanmeenen KM, McAuley JD, Beck KD, Pang KC, Servatius RJ. Behaviorally inhibited temperament is associated with severity of post-traumatic stress disorder symptoms and faster eyeblink conditioning in veterans. Stress. 2012; 15:31–44. [PubMed: 21790343]

North CS, Nixon SJ, Shariat S, Mallonee S, McMillen JC, Spitznagel EL, Smith EM. Psychiatric disorders among survivors of the Oklahoma City bombing. J Am Med Ass. 1999; 282:755–762.

North CS, Pfefferbaum B, Tivis L, Kawasaki A, Reddy C, Spitznagel EL. The course of posttraumatic stress disorder in a follow-up study of survivors of the Oklahoma City bombing. Ann Clin Psychiatry. 2004; 16:209–215. [PubMed: 15702569]

O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science. 2004; 304:452–454. [PubMed: 15087550]

O'Donnell ML, Elliott P, Lau W, Creamer M. PTSD symptom trajectories: from early to chronic response. Behav Res Ther. 2007; 45:601–606. [PubMed: 16712783]

O'Reilly RC, Frank MJ. Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. Neural Comput. 2006; 18:283–328. [PubMed: 16378516]

Parellada E, Lomena F, Catafau AM, Bernardo M, Font M, Fernandez-Egea E, Pavia J, Gutierrez F. Lack of sex differences in striatal dopamine D2 receptor binding in drug-naive schizophrenic patients: an IBZM-SPECT study. Psychiatry Res. 2004; 130:79–84. [PubMed: 14972370]

Pigott TA. Gender differences in the epidemiology and treatment of anxiety disorders. J Clin Psychiatry. 1999; 60(Suppl 18):4–15. [PubMed: 10487250]

Pohjalainen T, Rinne JO, Nagren K, Syvalahti E, Hietala J. Sex differences in the striatal dopamine D2 receptor binding characteristics in vivo. Am J Psychiatry. 1998; 155:768–773. [PubMed: 9619148]

Raia CP, Shillingford SW, Miller HL Jr, Baier PS. Interaction of procedural factors in human performance on yoked schedules. J Exp Anal Behav. 2000; 74:265–281. [PubMed: 11218225]

Schmajuk NA, Zanutto BS. Escape, avoidance, and imitation: A neural network approach. Adap Behav. 1997; 6:63–129.

Servatius RJ, Jiao X, Beck KD, Pang KC, Minor TR. Rapid avoidance acquisition in Wistar-Kyoto rats. Behav Brain Res. 2008; 192:191–197. [PubMed: 18501974]

Sheynin J, Beck KD, Pang KC, Servatius RJ, Shikari S, Ostovich J, Myers CE. Behaviourally inhibited temperament and female sex, two vulnerability factors for anxiety disorders, facilitate conditioned avoidance (also) in humans. Behav Processes. 2014a; 103:228–235. [PubMed: 24412263]

Sheynin J, Beck KD, Servatius RJ, Myers CE. Acquisition and extinction of human avoidance behavior: Attenuating effect of safety signals and associations with anxiety vulnerabilities. Front Behav Neurosci. 2014b; 15(8):323. [PubMed: 25309373]

Sheynin J, Shikari S, Gluck MA, Moustafa AA, Servatius RJ, Myers CE. Enhanced avoidance learning in behaviorally inhibited young men and women. Stress. 2013; 16:289–299. [PubMed: 23101990]

Smith A, Li M, Becker S, Kapur S. A model of antipsychotic action in conditioned avoidance: a computational approach. Neuropsychopharmacol. 2004; 29:1040–1049.

Stein MB, Paulus MP. Imbalance of approach and avoidance: the yin and yang of anxiety disorders. Biol Psychiatry. 2009; 66:1072–1074. [PubMed: 19944792]

Sutton, RS.; Barto, AG. Reinforcement learning: An introduction. Cambridge, MA: MIT Press; 1998.

Taylor JE, Sullman MJ. What does the Driving and Riding Avoidance Scale (DRAS) measure? J Anx Dis. 2009; 23:504–510.

Torrubia R, Avila C, Moltó J, Caseras X. The Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ) as a measure of Gray's anxiety and impulsivity dimensions. Pers Individ Dif. 2001; 31:837–862.

van Honk J, Schutter DJ, Hermans EJ, Putman P, Tuiten A, Koppeschaar H. Testosterone shifts the balance between sensitivity for punishment and reward in healthy young women. Psychoneuroendocrinology. 2004; 29:937–943. [PubMed: 15177710]

Highlights

1. A reinforcement-learning model successfully simulated human avoidance behavior.
2. Distinct reward and punishment sensitivity ratio might underlie sex differences.
3. Distinct punishment sensitivity might underlie inhibited temperament differences.
4. Attenuating effect of safety-signals is due to the competing approach response.
5. Safety-signals might be used in cognitive-behavior therapies to reduce avoidance.

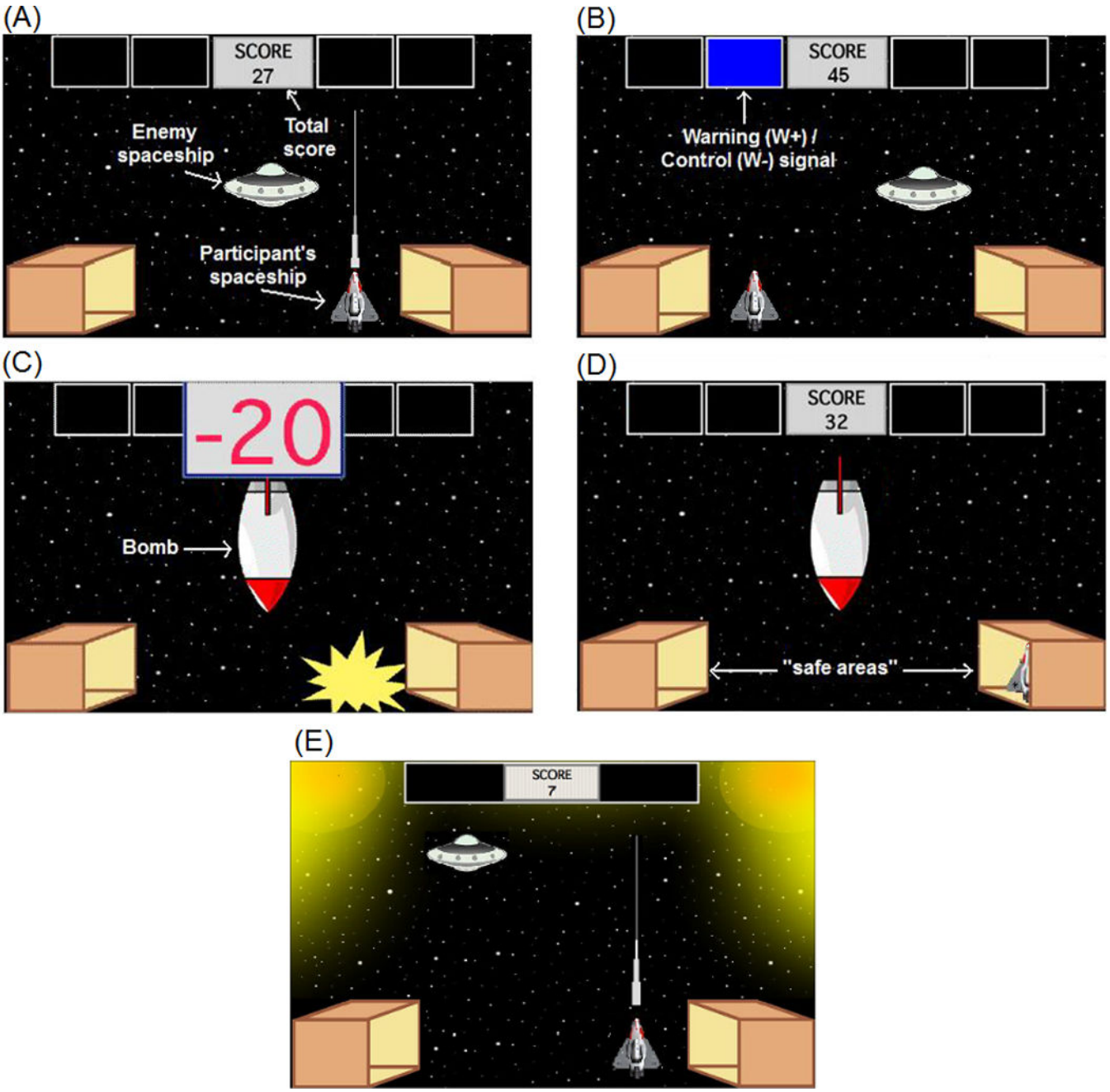

Fig. 1.

Computer-based escape-avoidance task (adapted from Sheynin et al. 2014a (A–D); Sheynin et al. 2014b (E)). (A) Participants controlled a spaceship located at the bottom of the screen and were instructed to maximize their score. They could learn that a reward (one point) could be obtained by shooting and destroying an enemy spaceship that was moving on the screen. (B) In the original study (Sheynin et al. 2014a), every 20 s, a signal (a colored rectangle at the top of the screen) appeared for 5 s. Depending on its color, the signal could be a warning signal (W+) or a control signal (W−). (C) W+ was always followed by appearance of a bomb, which remained onscreen for another 5 s (bomb period), during

which a maximum of 30 points could be lost. (D) Participants could learn to protect themselves from this aversive event by moving their spaceship to a specific “safe area” on the screen; moving there during the bomb period terminated the point loss (ER), while moving during the warning period could completely prevent any point loss (AR). Importantly, while in the safe area, it was impossible to shoot the enemy spaceship and obtain reward. (E) Sheynin et al. (2014b) modified this task by eliminating control trials and adding an SS (yellow lights) which appeared during the ITI on acquisition trials. Labels shown in white text are for illustration only and did not appear on the screen during the task.

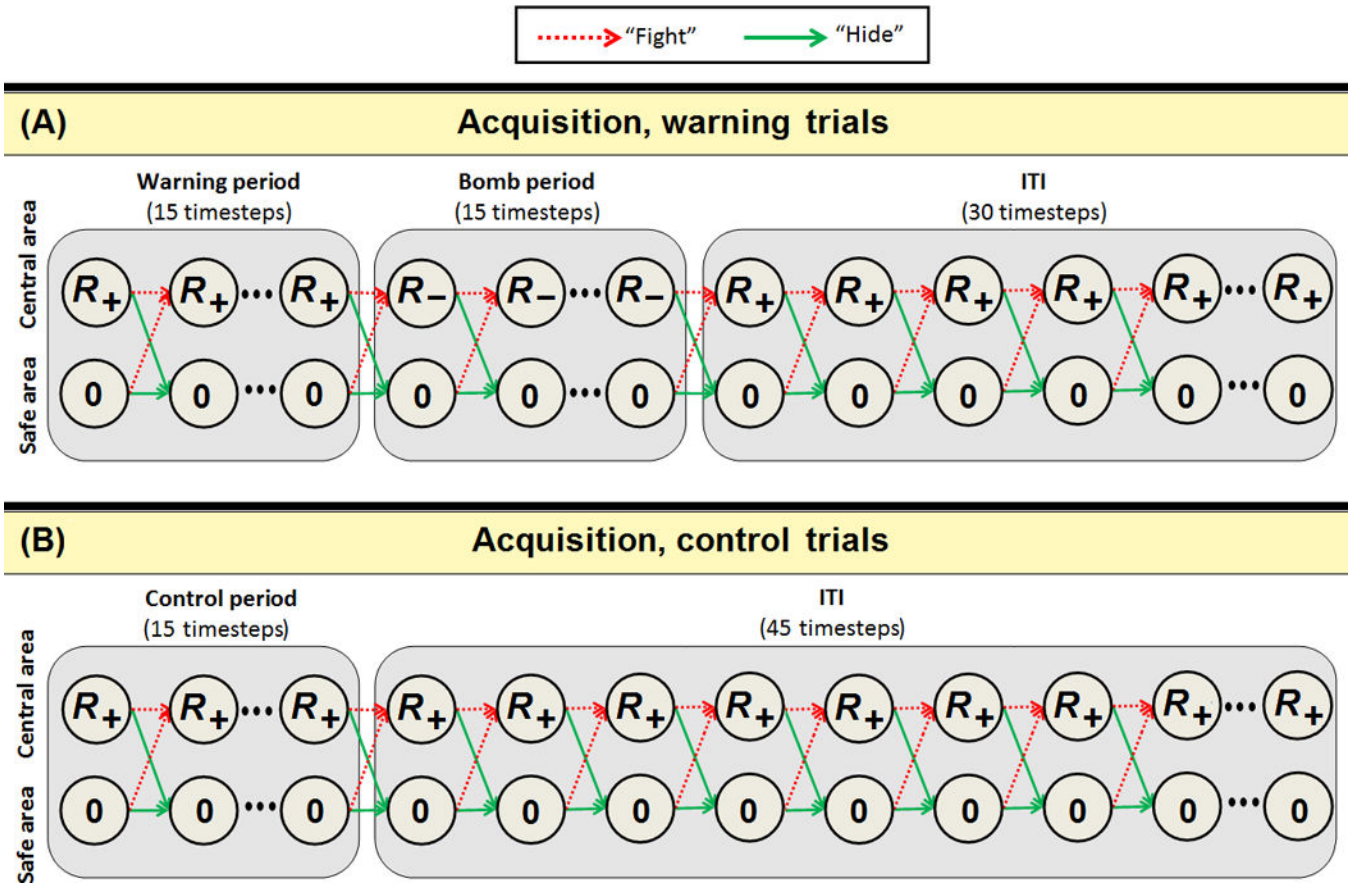

Fig. 2.

Decision process representation of a trial during the acquisition phase of the computer-based avoidance paradigm. Each simulated trial was divided into 60 timesteps, where each timestep represented 0.33 s. On each timestep, a total of four inputs were provided to the model: presence or absence of a warning signal (W+), a control signal (W−), an SS and an aversive event (bomb). On each timestep, the agent chose between two actions – "fight" (remain/move to the central area; dashed red arrows) or "hide" (remain/move into a "safe area"; solid green arrows). When located in the central area during the bomb period, external reinforcement was R_{punish} (point loss; depicted by R_{-}). When located in the central area during the warning period or the ITI, reinforcement was R_{reward} (point gain; depicted by R_{+}). When located in a "safe area", reinforcement was always zero (no points gained or lost). (A) Warning trials included 15 timesteps with W+, followed by 15 timesteps with a bomb, followed by 30 timesteps of ITI. (B) Control trials included 15 timesteps with W−, followed by 45 timesteps of ITI (no bomb period).

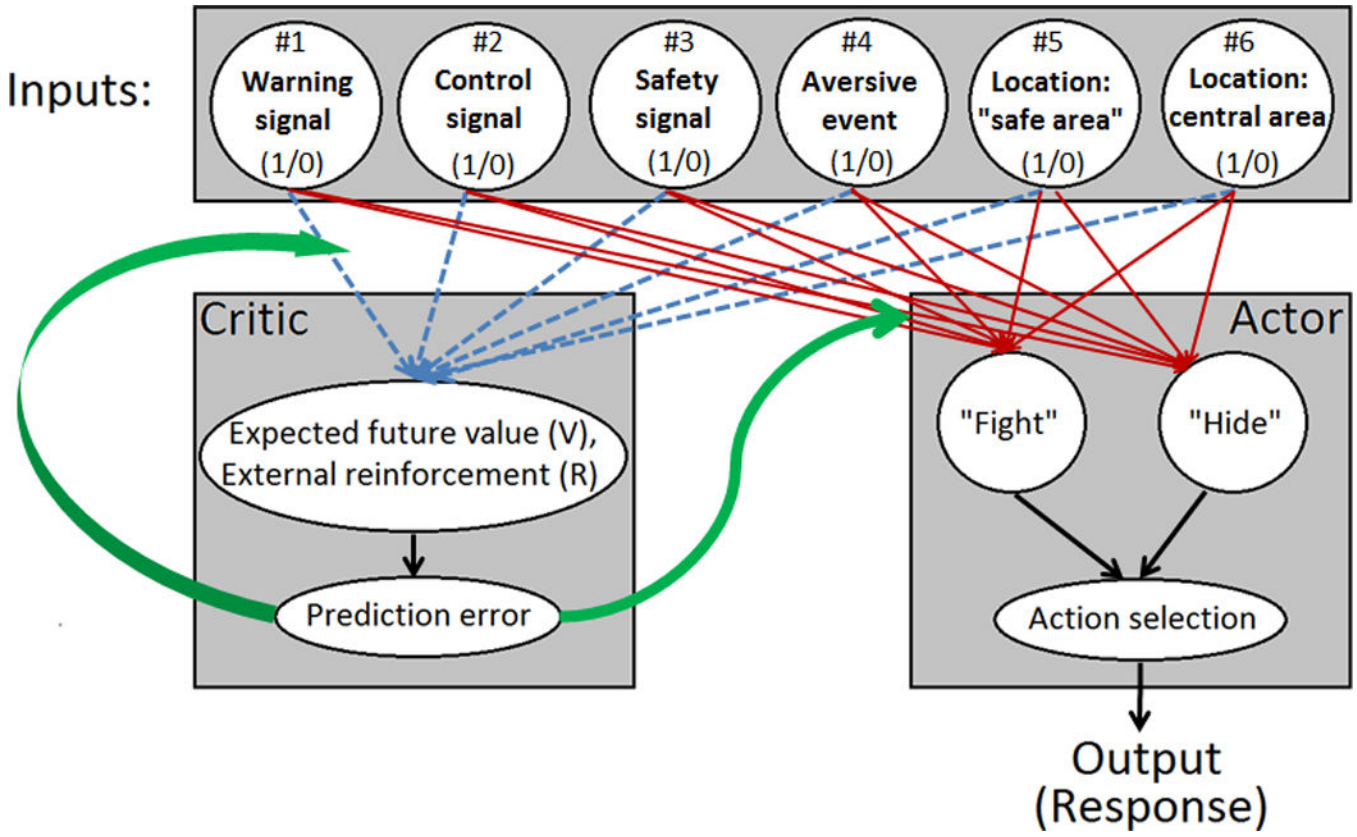

Fig. 3.

The network model (“agent”) architecture. On each timestep, six binary inputs signaled the presence or absence of the (1) warning signal (W+) (2) control signal (W−) (3) safety signal (4) aversive event (bomb), as well as the location of the subject’s spaceship – (5) inside a “safe area” or (6) in the central area of the screen. On each timestep, the critic computed prediction error (the change in expected future reinforcement; Eq. 1). This prediction error was used to update (thick green arrows) the connections from the inputs to the critic (Eq. 3; dashed blue arrows), as well as the connections from the inputs to the actor (Eq. 6; solid red arrows). These updated connections in the actor were used to generate a response, according to a softmax probabilistic rule (Eq. 4).

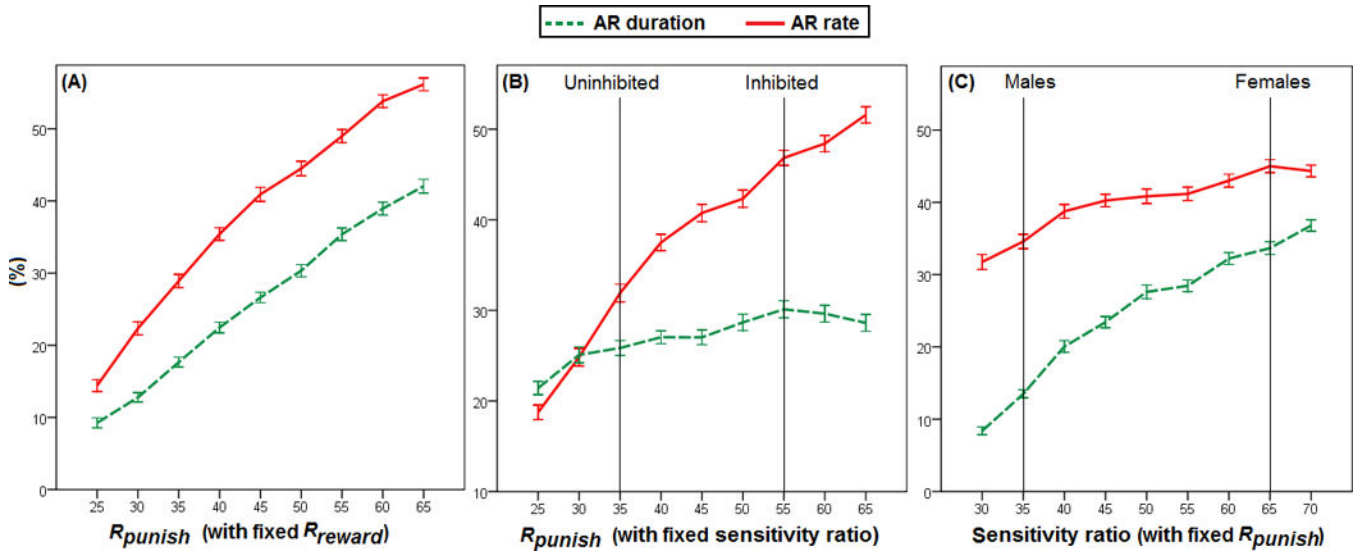

Fig. 4.

AR as a function of parameter manipulations in the model. (A) AR rate and AR duration (solid red and dashed green lines, respectively) as a function of increasing sensitivity to punishment but fixed sensitivity to reward ($R_{reward} = -0.9$). (B) AR rate and AR duration as a function of increasing sensitivity to punishment but fixed sensitivity ratio (ratio=50; i.e., sensitivity to reward was proportionally increased). (C) AR rate and AR duration as a function of increasing sensitivity ratio but fixed sensitivity to punishment ($R_{punish} = 45$). Based on these results, values were set to define "females" versus "males" [vertical lines in (C)] and "inhibited" versus "uninhibited" [vertical lines in (B)] simulations, as described in Fig. 5. Error bars indicate SEM.

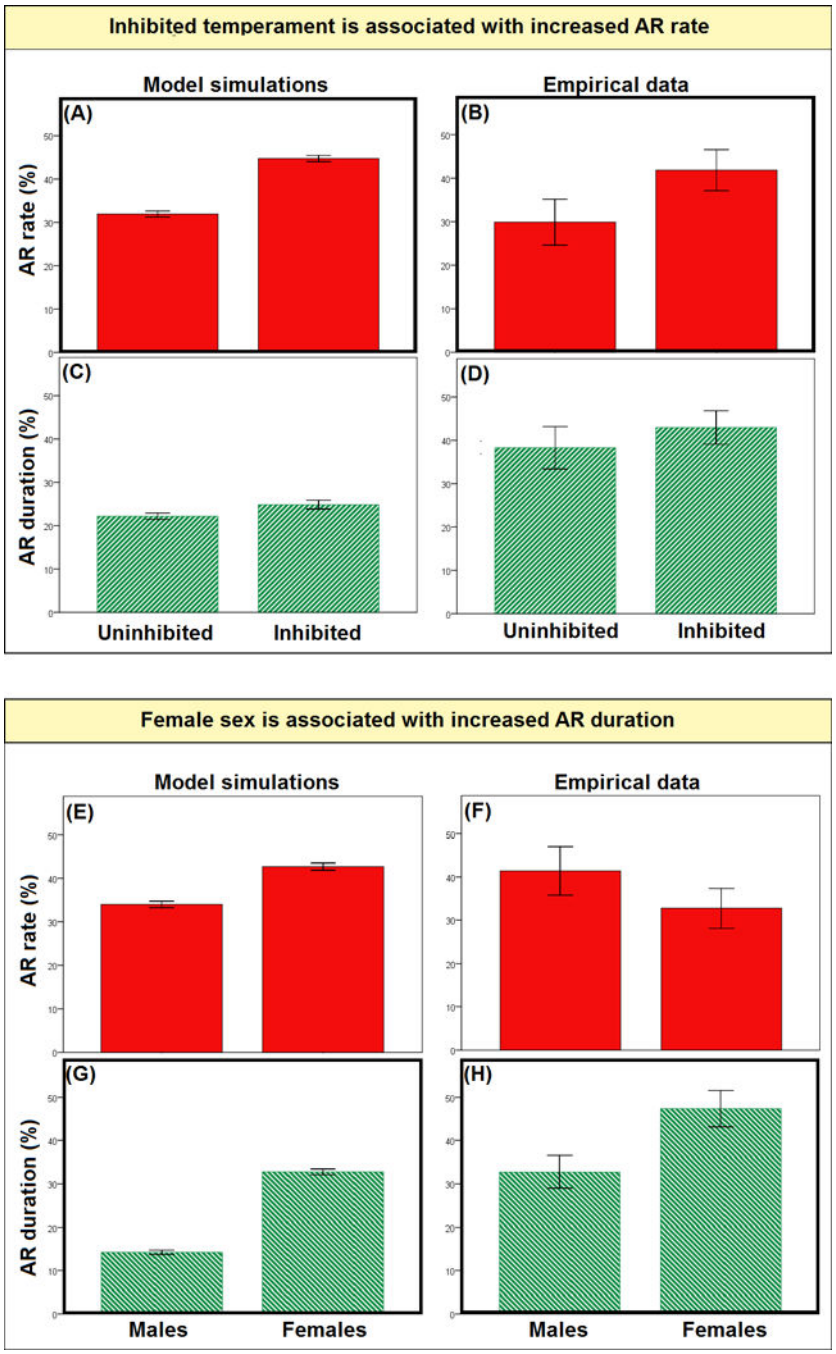

Fig. 5.

AR rate and AR duration values (solid red and striped green, respectively) of inhibited versus uninhibited subjects (A–D) and of males versus females (E–H), in the model (left) and empirical data (adapted from Sheynin et al. 2014a). Both the model and the empirical data suggest that inhibited temperament is a stronger predictor of AR rate (A–D), whereas sex is a stronger predictor of AR duration (E–H). See the main text for a detailed explanation of these analyses, as well as a discussion on the discrepancy between simulated

and empirical findings on AR rate in males and females (Fig. 5E,F). Error bars indicate SEM.

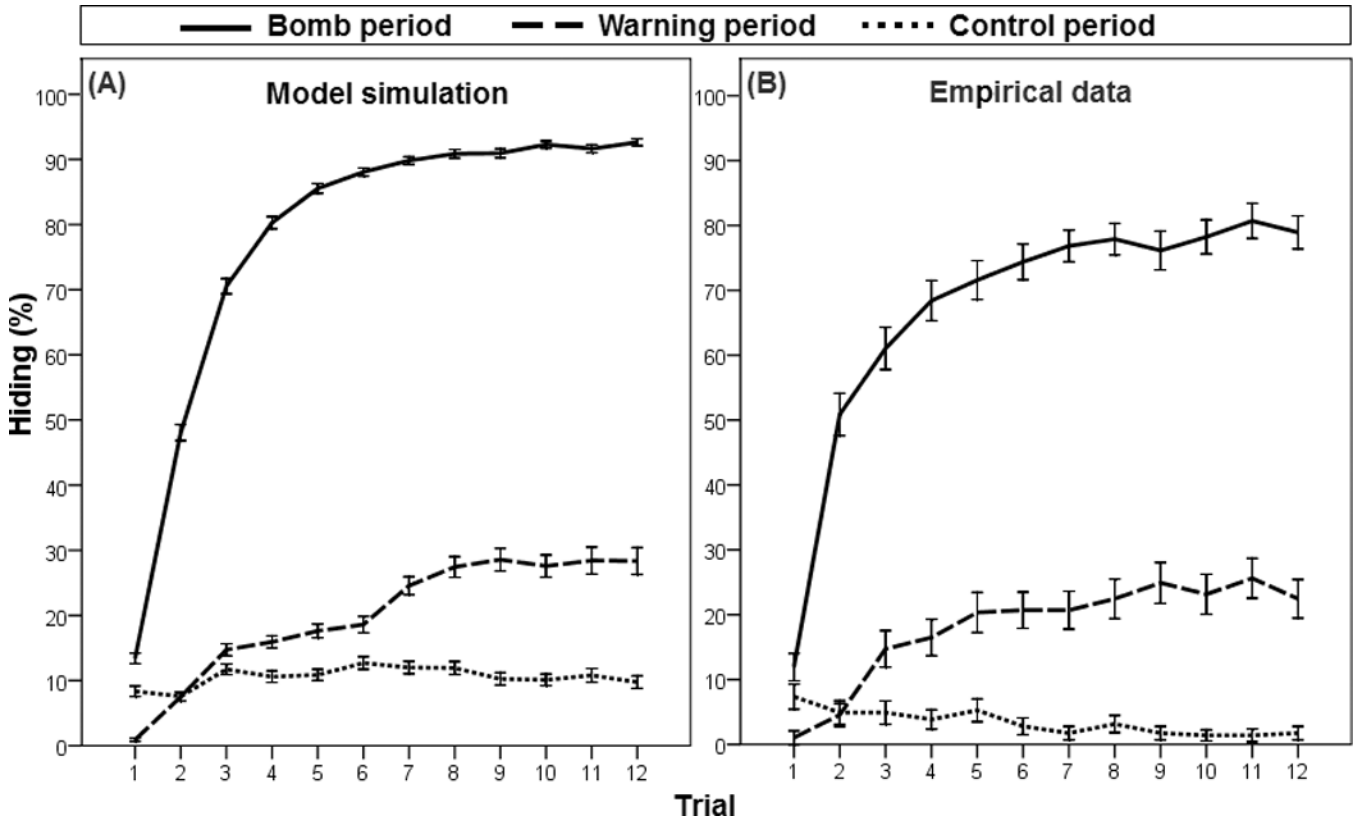

Fig. 6.

Total hiding during the different task periods (dotted = control period; dashed = warning period; solid = bomb period), in (A) the model and (B) empirical data (adapted from Sheynin et al. 2014a). Both the model and the empirical data show that subjects quickly learn the ER, gradually learn the avoidance behavior and show little hiding during the control period. Model simulations were run the same number of times as the different group sizes in the empirical data [i.e., uninhibited males (n=22), uninhibited females (n=24), inhibited males (n=14) and inhibited females (n=35)]. Error bars indicate SEM.

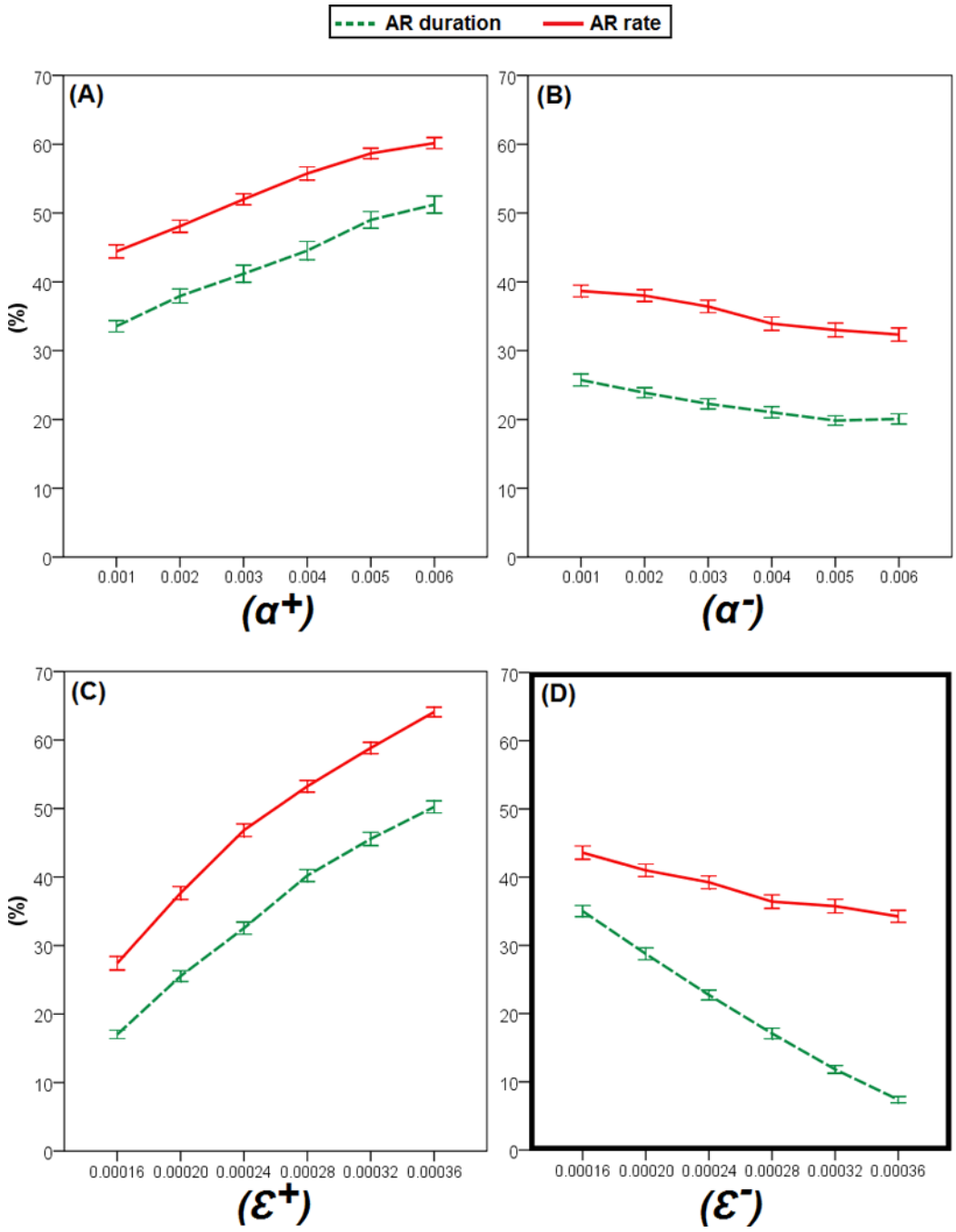

Fig. 7.

AR rate and AR duration (solid red and dashed green lines, respectively) as a function of specific LR manipulations in the model. (A) Manipulating the LR associated with positive PEs in the critic; α^{+}. (B) Manipulating the LR associated with negative PEs in the critic; α^{-}. (C) Manipulating the LR associated with positive PEs in the actor; ε^{+}. (D) Manipulating the LR associated with negative PEs in the actor; ε^{-}. Overall, results support the idea that changes in LRs that are associated with negative PEs might parallel the sex differences in AR duration described in the current study. Error bars indicate SEM.

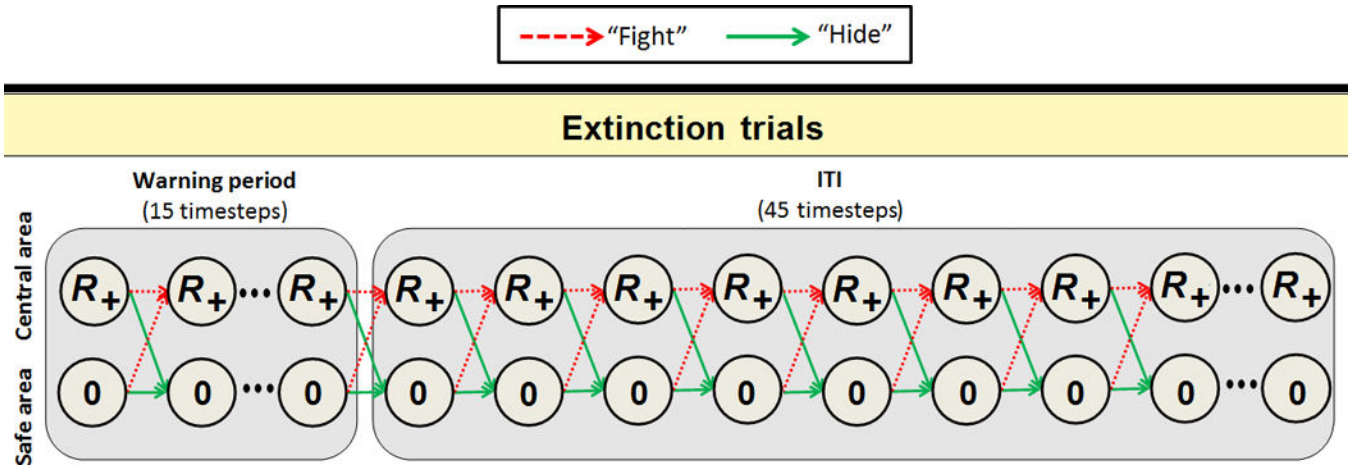

Fig. 8.

Decision process representation of a trial during the extinction phase in the computer-based avoidance paradigm. All extinction trials included 15 timesteps with a W+ signal, followed by 45 timesteps of ITI. The acquisition phase followed a similar design as described earlier, but only including warning trials (Fig. 2A).

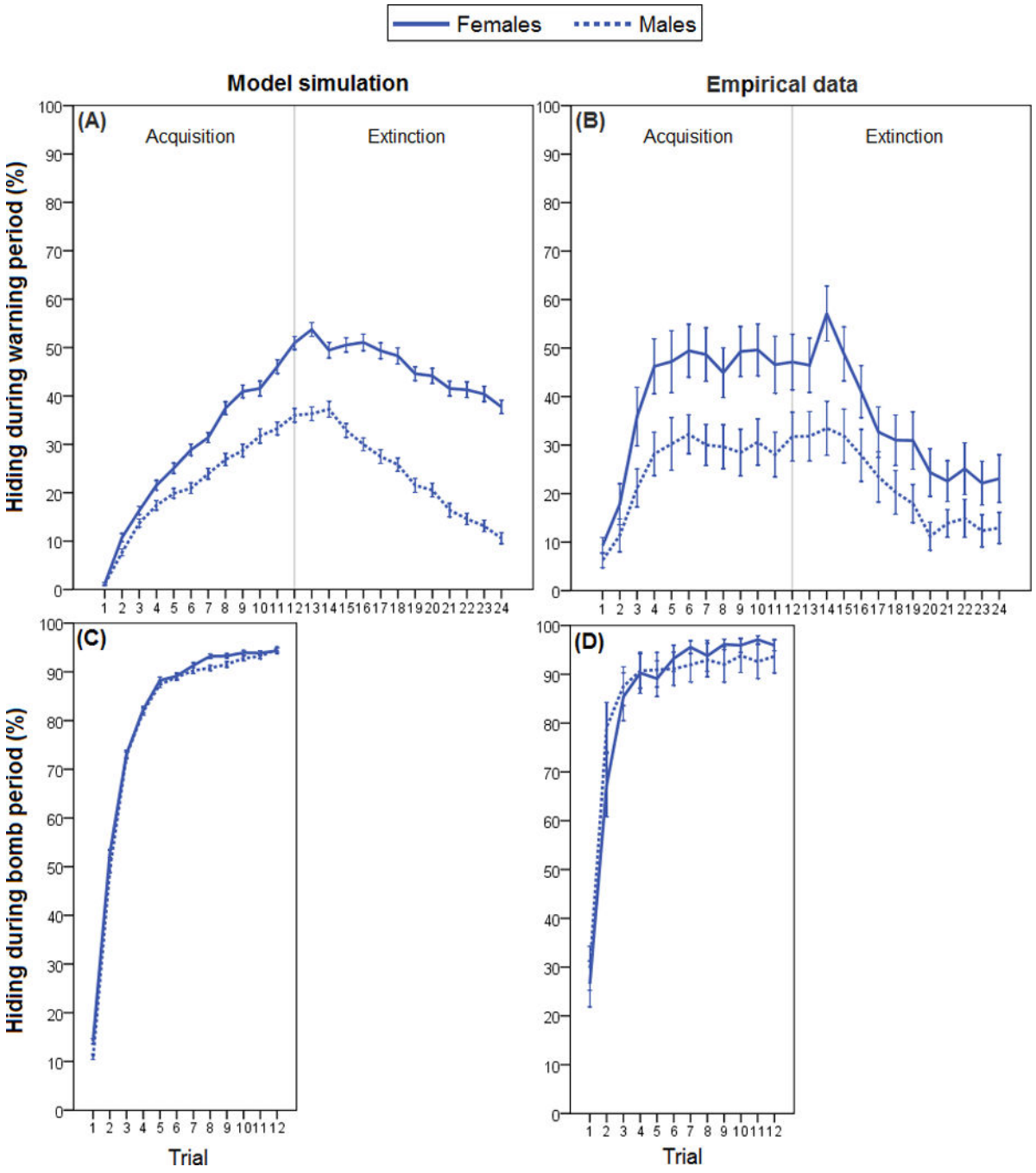

Fig. 9.

Hiding during the warning period [(A–B) avoidance behavior] and during the bomb period [(C–D) ER] during acquisition (trials 1–12) and extinction (trials 13–24). Note that there is no bomb period during extinction trials. Analyses compared hiding performance in male versus female individuals (dotted versus solid lines, respectively). In both the model (A,C) and empirical data (B,D; adapted from Sheynin et al. 2014b “without-SS” group), females showed facilitated acquisition and slower extinction of the avoidance behavior compared to males, with no difference on ER. Error bars indicate SEM.

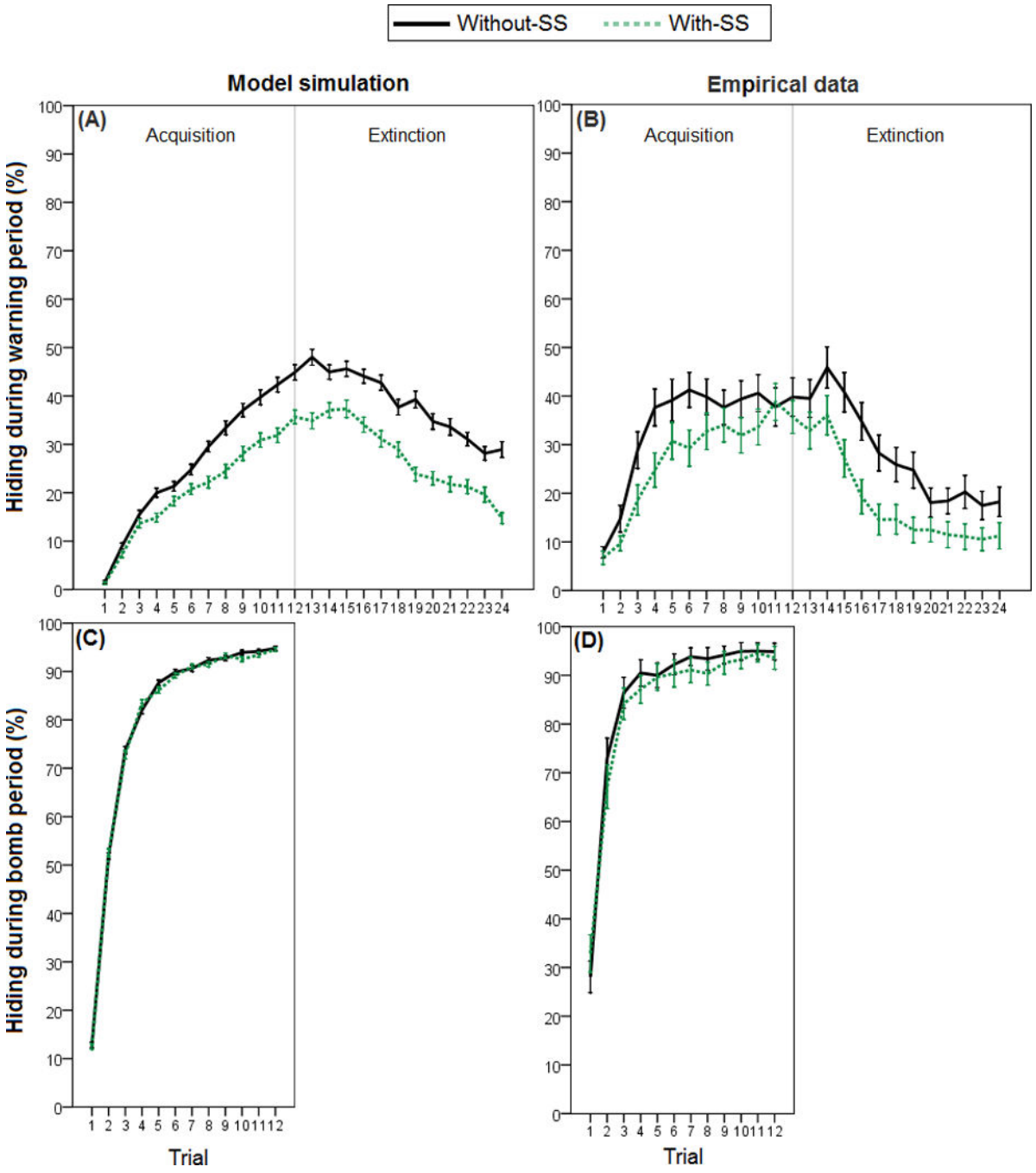

Fig. 10.

Hiding during the warning period [(A–B) avoidance behavior] and during the bomb period [(C–D) ER] for acquisition trials 1–12 and extinction trials 13–24 (note that there is no bomb period in extinction trials). Analyses compared the behavior of subjects performing an avoidance task with versus without SS (dotted green versus solid black lines, respectively). In both the model (A,C) and empirical data (B,D; adapted from Sheynin et al. 2014b), the SS facilitated the extinction of the avoidance behavior, without affecting ER. Error bars indicate SEM.

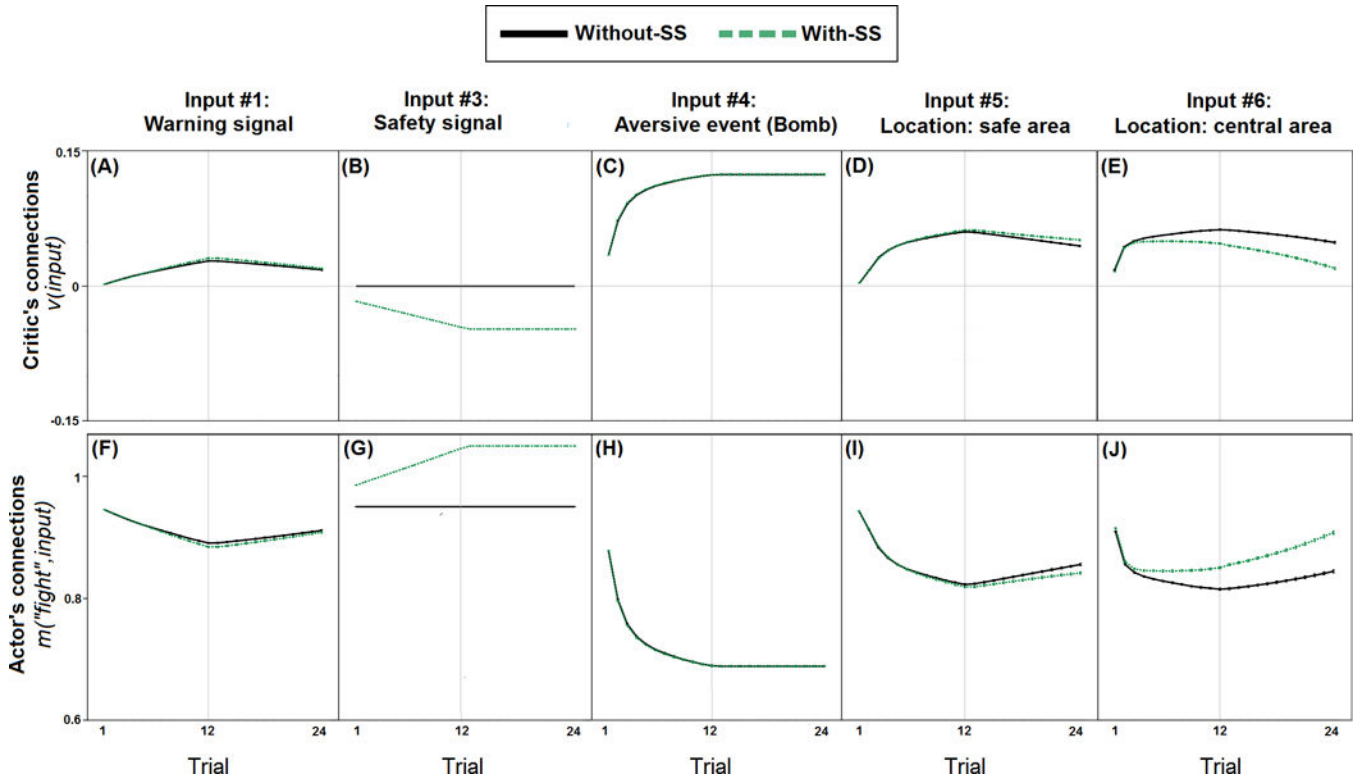

Fig. 11.

Changes in model weights across 12 acquisition and 12 extinction trials, with and without the presence of an SS during ITI (and "practice" period) on acquisition trials (dotted green versus solid black lines, respectively). Overall, data suggest that simulated behavioral changes were driven mainly by decreased values (decreased predicted punishment) of the connections associated with the presence of the SS and the state of being at the central area (B,E), which in turn, resulted in increased probability to choose the "fight" response when these inputs were activated (G,J). For a detailed explanation of these analyses, see main text. For clarity, several features were omitted from the figure: input #2 was omitted as it represents the presence of the control signal, which was always set to zero in this simulation; the actor's connections associated with the "hide" response were omitted due to a minimal change in their values for all the inputs (range: 0.8493–0.8559) – suggesting that learning affected mainly the competing "fight" response. Error bars indicate SEM.

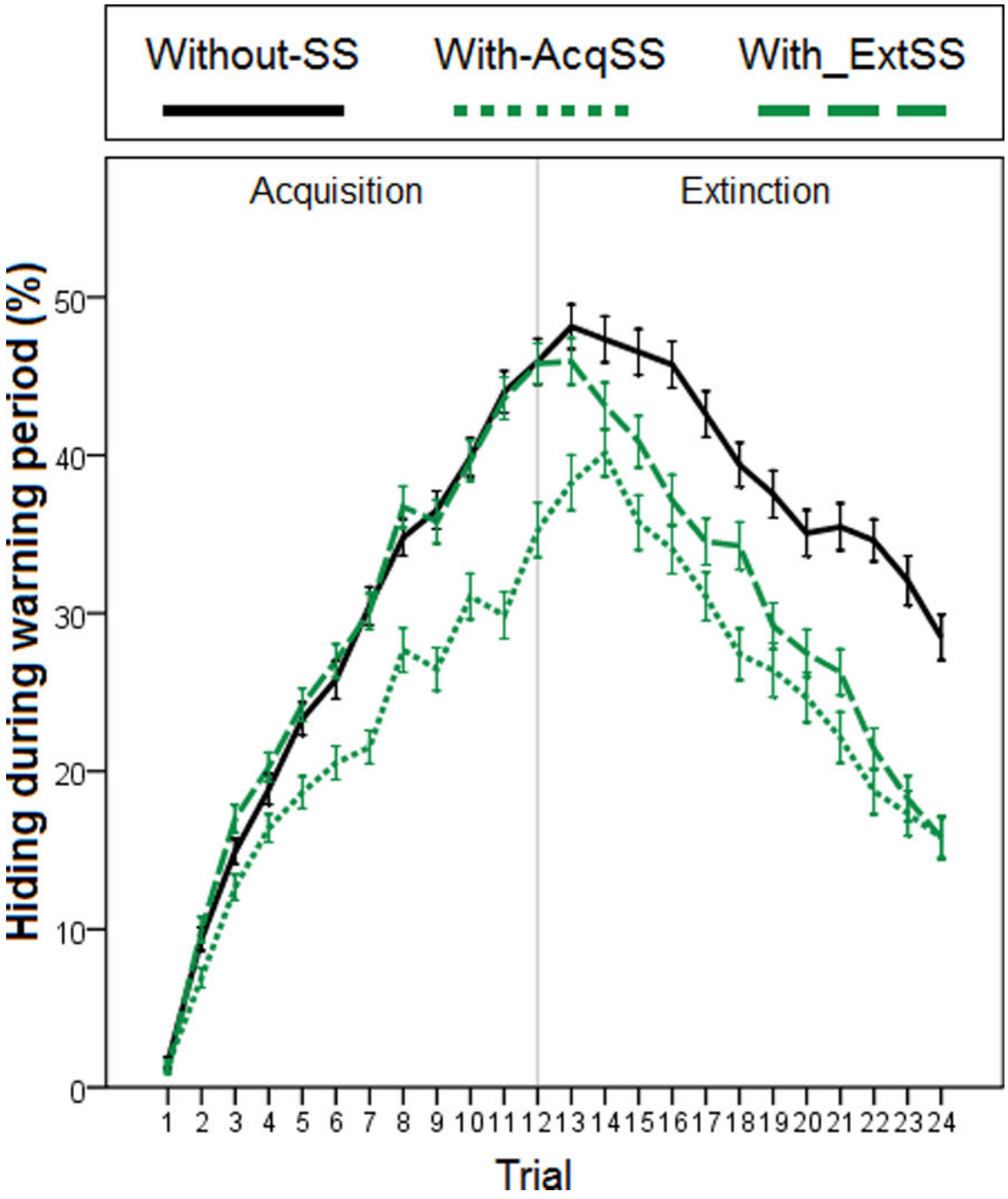

Fig. 12.

Simulating the effect of an SS administered during different phases of the task. An SS administered during the acquisition phase (dotted green line) or extinction phase (dashed green line) facilitated extinction learning, compared to simulations without an SS on either phase (solid black). Error bars indicate SEM.

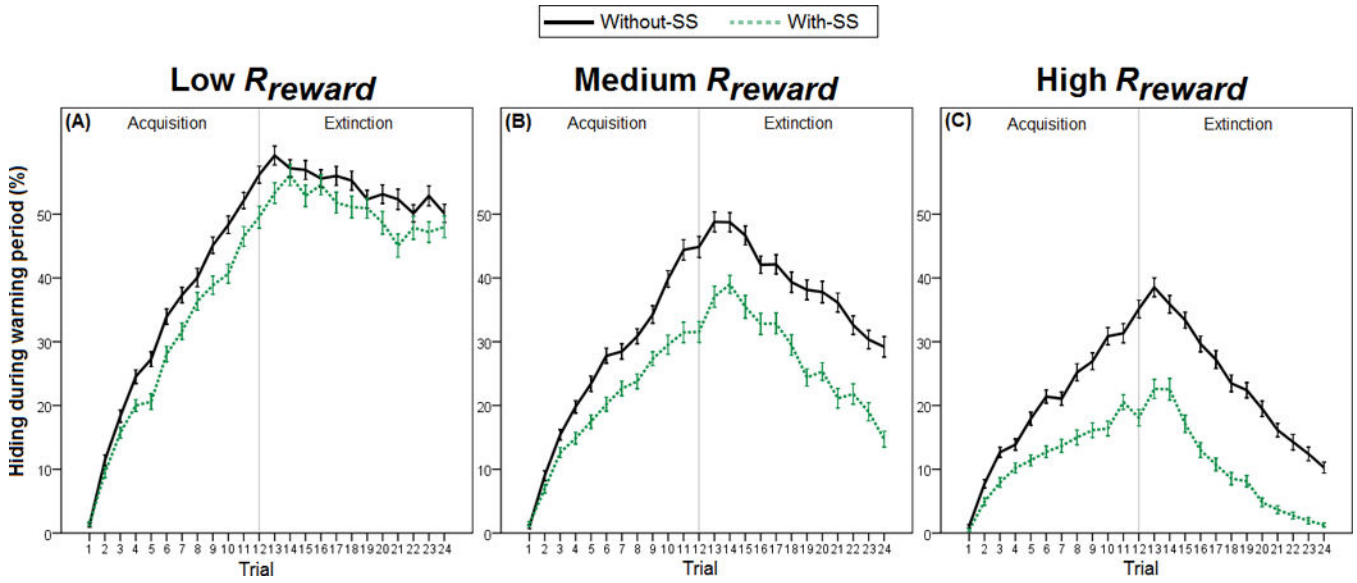

Fig. 13.

Simulating the effect of SS administration in models with different reward (but fixed punishment) sensitivity values [(A–C) R_{punish} = 45 (A) low R_{reward} = −0.45 (B) medium R_{reward} = −0.9 (C) high R_{reward} = −1.29]; “with-SS” (green dotted line) versus “without-SS” (black solid line). While all simulations show an attenuating effect of the SS on avoidance, this effect was stronger in simulations with higher R_{reward}. Error bars indicate SEM.

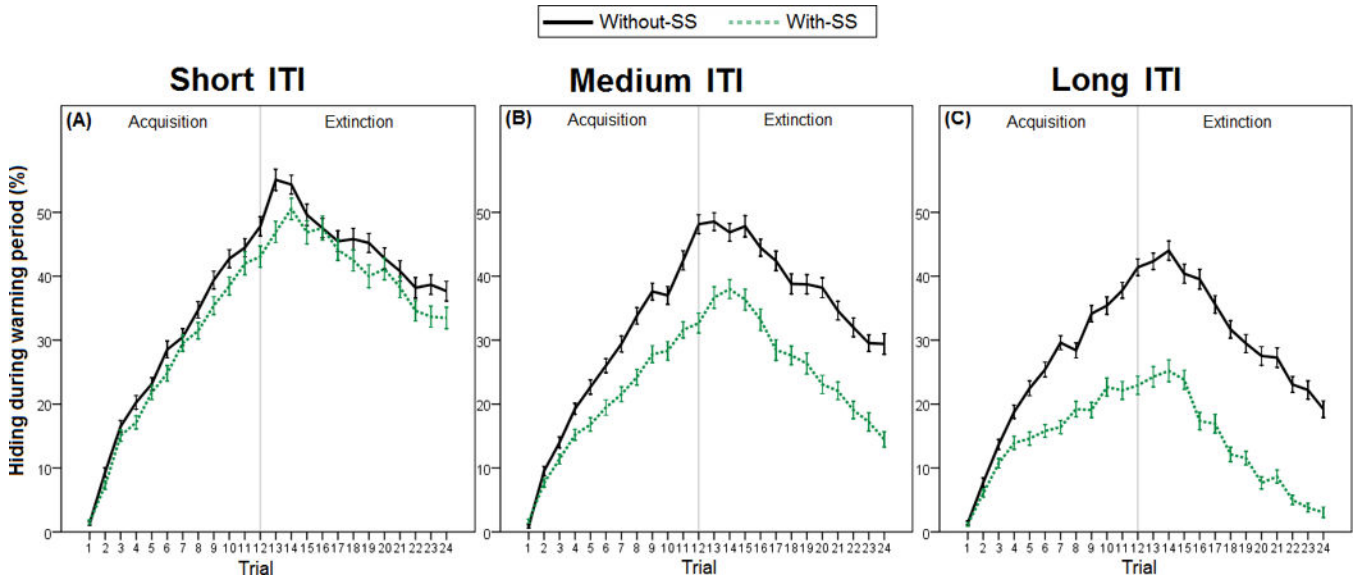

Fig. 14.

Simulating the effect of SS administration during the ITI on acquisition trials, in models with different ITI durations [(A) short ITI = 10 timesteps (B) medium ITI = 30 timesteps (C) long ITI = 50 timesteps]; “with-SS” (green dotted line) versus “without-SS” (black solid line). While all simulations show an attenuating effect of SS on avoidance, this effect was dependent on the ITI duration and was stronger when ITI was longer. Error bars indicate SEM.