Supplementary Materials for

Toward quantitative cognitive-behavioral modeling of psychopathology: An active inference account of social anxiety disorder

Table of Content

Supplementray Appendix 1: Model specification

The computational model was endowed with the structure of partially observable Markov decision process (POMDP) and built under the active inference framework. Please refer to the main text for theoretical rationale and (Parr et al., 2022) for mathematical formalizations of active inference. Here, we elaborate on each model component. This section is best understood with the MATLAB code.

First, we specify the parameter values for four free parameters (*PriorSafe, SEff, PostRum, SAttn*) to simulate agents with different characteristics in the beginning (Lines 9-20). In the supplementary code, these parameters were set to simulate a healthy agent. However, one can try different values for each parameter to simulate various agent with different SAD traits and symptom severity.

Prior beliefs about two sets of hidden state factors were specified as probability distributions in $\boldsymbol{D}$ vectors. $\boldsymbol{D_1}$ and $\boldsymbol{d_1}$ matrices encoded prior beliefs about the hidden action states, that the agent always started in the 'start' state ([1 0 0 0 0 0 0], each data point correspond to a potential action state in the order of 'start', 'attend to environment', 'attend to self', 'stay', 'escape', 'realistic reflection', 'rumination'; lines 28-29, line 42). Prior beliefs about the social context states were specified in $\boldsymbol{D_2}$ and $\boldsymbol{d_2}$ matrices (lines 32 and 40), where the true social context was always more threatening in the first 30 trials (for N=1:30, $\boldsymbol{D_2}$ = [.1 .9] ([p(safe) p(threatening)]); line 268) and safer in the subsequent trials (for N=31:100, $\boldsymbol{D_2}$ = [.9 .1]; line 274). All agents, except for the agent with multiple vulnerabilities to SAD, entered the simulations with an optimistic bias towards the social context state ($\boldsymbol{d_2}$ = [*PriorSafe* 1-*PriorSafe*], where *PriorSafe* is a free parameter representing the agent's prior beliefs about social context). This belief was updated during the 100-trials simulation, representing context learning. Learning in active inference models involves adding counts to Dirichlet parameters coded in the lower-case vectors/matricesUpdated belief (i.e. the posterior) of the n^{th} trial acted as the prior belief for the $(n+1)^{th}$ trial.

A matrices encoded beliefs about the likelihood mappings between hidden state factors and observable outcomes. Rows in each ***A*** matrix corresponded to possible outcome observations of each type, while columns are indexed by the seven action states (‘start’, ‘attend to environment’, ‘attend to self’, ‘stay’, ‘escape’, ‘realistic reflection’, ‘rumination’). The third dimension of ***A*** matrices distinguished two possible social context states. The probabilities of observing different Interoceptive outcomes, namely two arousal levels, were specified in A_1 (lines 74-83). Agents were very likely to observe high arousal levels if they had heightened self-focused attention (3^{rd} coloum) or ruminated (last column). Besides, agents were slightly more likely to observe high arousal if they stayed in the social interaction when the context was threatening. The likelihood of observing three possible exteroceptive observations, namely judgemental, friendly or neutral facial expressions of surrounding people were specified in A_2 in a similar fashion (lines 85-97). The precision of $\boldsymbol{A_2}$ matrices was significantly higher than $\boldsymbol{A_1}$**,** corresponding to the fact that external information was more indicative of social context. $\boldsymbol{A_3}$ encoded the probability of observing three possile social consequences (‘mocked’, ‘neutral’, ‘respected’) perceived by the agent (lines102-110). The agent only perceived social consequences during the approach/avoid and reflection stages, and the probability of being respected in social encounters was dependent on the agent’s social self-efficacy. With the level of self-efficacy being identical, the agent was five times more likely to observe success social interations where they were respected in safe than threatening social contexts. Lastly, $\boldsymbol{A_4}$ was a set of identity matrice specifying that the agent could always 100% accurately perceive which action state they were in (lines 113-119).

The transition probability between hidden states within a simulation trial were encoded by ***B*** matrices. Specifically, the social context state remained stable for the agent within a given trial, such that no transition between different social context states was allowed (lines 124-126). On the other hand, the agent was freely able to transit between action states (e.g. from 'start' to 'attend to self') through policy selection. These action-dependent state transitions were specified in six ***B*** matrices expanded in the code. Each ***B*** matrix corresponded to one possible action at a transition point, yielding two possible actions for each of the three transition as annotated in the code (lines 128-174). No action stochasticity was involved for state transitions, such that the agent always had a probability of 1 to transit to the state according to their intention. Together, these actions produced eight different policies that were stored in the ***V*** matrix, where each row specied a different policy composing of three actions (lines 185-195). It was specified in ***E*** vector that the agent did not start with any pre-existing preference for choosing any of these policies ([1 1 1 1 1 1 1 1], a flat probability distribution ; line 200). The agent was able to learn and form habits over trials, with the updated information stored in the ***e*** vector (lines 202-203), such that they could develop prior preference for certain policies that have been most frequently chosen. After trying out several different values, we decided to multiply ***e*** by 512 to control the effect of habit learning within a reasonable range.

The agent's prior preference for observing certain sensory outcomes was specified in ***C*** vectors. The way positive and negative values specified in ***C*** vectors work resemble a reward/punishment system. Positive values in larger magnitude were converted to a higher prior probability of observing those corresponding sensory outcomes by enactively taking actions to enter the states that were most likely to generate these outcome observations. Similarly, negative values indicates the agent's aversion for certain sensory outcomes and reduces their prior expectations for encountering those observations. The agent had no prior preference for most of the outcome observations (i.e. the reward/punishment values were 0, meaning that most of the observations were perceived neutrually, lines 208-211) except for those specified in particular. Here, negative interoceptive and exteroceptive sensory observations, namely high arousal and judgemental facial expression were slightly ‘punishing’ for the agent, while positive (being respected) and negative (being mocked) social consequences carried large reward/punishment values and therefore relatively larger influences on the agent’s policy selection (lines 213-216). Lastly, the agent’s prior preferences for observing themselves in states of ‘attend to self’ and ‘rumination’ were specified by free parameters that were manipulatable on the top of the script (lines 218-219). By tuning these two parameters, one can simulate agents with different levels of tendency for inwardly focused attention and rumination.

Supplementary Appendix 2: A step-by-step simulation walk through for healthy agent

At timepoint 1 ($\tau = 1$) of the 1st trial, the agent's action state was at "start". The agent observed baseline sensory outcomes (Figure S2, observation panel). The agent was then tasked to choose an action that would determine attention allocation in the next time point. As aforementioned, when attending to self, the agent would be more likely to observe a negative sensory outcome (e.g., "high" arousal) which was undesirable. Besides, "attend to environment" would yield more information than "attend to self" about the social context. This means that the action "attend to environment" had both higher utility and epistemic values than "attend to self". Consequently, policies including "attend to environment" would have higher probabilities than policies including "attend to self" for the healthy agent, increasing the probability of selecting the action of "attend to environment" (Figure S2, action panel). Based on this probability distribution, an action was drawn at the end of timepoint 1, which then brought the agent to a new action state at timepoint 2.

At $\tau = 2$, the healthy agent would again sample the observable outcomes and decide what action to take. Given the true social context was more likely to be threatening, the agent who was attending to the environment was likely to observe negative exteroceptive information. This new piece of sensory evidence would then help update the agent's belief about the social context, informing action selection at the next transition point. To choose between "stay" and "escape", the agent needed to predict the observable outcomes in the future time points, including interoception, exteroception, and importantly social consequences for each of the two action states. If the healthy agent inferred that the social context was threatening, they would have a higher probability of escaping to prevent social loss. However, in the first few trials, with the slightly optimistic prior belief about the social context, a healthy agent might still choose to stay in the social context to gain the potential social reward from successful social interactions.

Moving to $\tau = 3$, the agent was likely to observe a series of negative sensory outcomes given the threatening context if they stayed. However, the perceived social consequence could still be positive, which would be attributed to the high level of self-efficacy. If the agent chose to escape, they would gain little information about the true social context. The conflict between wanting to gain information as well as potential social reward and wanting to avoid potential social loss was prominent here. This highlights that the agent needed to take the action that could minimize free energy which balances the epistemic value and utility (Friston et al., 2015).

Finally, at $\tau = 4$, a healthy agent without preferences for "rumination" would always transit to the state of realistic "reflection", because if they chose "rumination" then undesirable sensory outcomes were likely to be perceived. Here, in "reflection" state, the agent observed sensory observations that were consistent with the true social context. This provided further information about the social context. Note that at the initialization phase of a new trial, the prior belief about the social context and policy selection preference were updated (values in $\boldsymbol{d}$ and $\boldsymbol{e}$) given the experiences gained from the previous trial.

Supplementary Simulation 1: Simulation for agent with altered learning rates

As stated in the maintext, we emphasize the importance of the relationship between learning rate and SAD psychopathology by including an additional simulation here. Healthy individuals have been found to exhibit an elevated learning rate for positive social consequences (Koban et al., 2017). Conversely, individuals with SAD tend to possess an increased learning rate for negative outcomes, which is also commonly observed in other internalizing disorders such as generalized anxiety disorder and depression (Pike & Robinson, 2022). As a result, negative observations during the social encounter, including interoceptive and exteroceptive cues as well as negative social consequences would have a stronger impact on these individuals' belief updating, promoting safety behaviors. We simulated this factor by endowing the agent with higher learning rate (η=1) when the true social context was threatening (i.e. in the first 30 trials) than when the true context was safe (η=.2, in the subsequent 70 trials), with all other parameters set as the same values as the healthy agent (**Table 1** in main text). Note that under the active inference framework, "learning" refers to how the agent update their beliefs about the most optimal policy after each trial. To magnify the effects of altered learning, the precision of this agent's prior preferences about policies and prior belief about the social context were lowered, so that the agent would have faster habit learning and belief update about the social context. For this agent, we predicted that maladaptive policies would be more difficult to change once they were acquired.

With a high learning rate in negative environment, the agent quickly learned to escape from social encounters after a single social mishap and sticked to this action for the remaining of the first 30 trials (**Figure S1A**). This pattern demonstrated a strong avoidance acquisition in a hostile environment. Importantly, once a habit is formed it is harder to change. Under active inference, this is reflected by smaller free energy for habitual (highly expected) options (Smith et al., 2022). Consequently, avoidance tended to persist. In the later 70 trials, the agent experienced positive social consequences when they chose to stay. However, because of the lower learning rate in the positive environment relative to the learning rate in the negative environment, these positive experiences were not able to alter the action substantially.

Results from this simulation resemble the clinical presentation marked by prolonged avoidance following a major adverse social event. The simulation provided possible explanatory process of why some patients fail to adjust their learned behaviors, showing rigidity in policy selection, even after correcting their cognitive distortion (Caletti et al., 2022; Hayes et al., 2006). When our active inference agent had a higher learning rate in a hostile environment compared to a supportive one, they quickly acquired an avoidant policy that became habitual, even after experiencing a single social mishap. This avoidant habit can contribute to a reduction in behavioral flexibility, prolonging SAD symptoms (Alvares et al., 2014). Additionally, a relatively lower learning rate in a positive environment can result in reduced sensitivity to positive social outcomes. Surprisingly, the simulated agent did not develop inaccurate beliefs about the social context, despite the strong effect on habit learning. This suggests that inflexible actions can arise from habit learning that is relatively unaffected by the individual's beliefs about the environment (Gillan et al., 2016). The current results also suggest that individuals who are more sensitive to threat than safety cues (i.e., those with a larger learning rate in threatening environments) may require more exposure sessions.

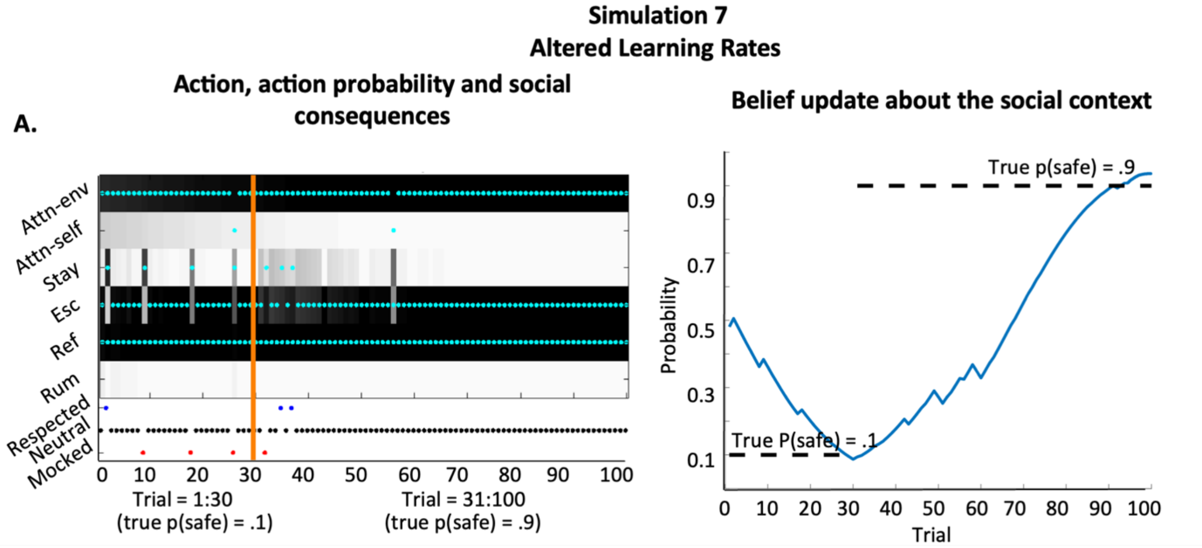
Supplementary Simulation 2: Simulation for agent with two changes in the environment

In this simulation, a healthy agent and an agent with lower self-efficacy are exposed to three distinct environments. The first 30 trials occurred in a predominantly safe environment, with .9 probability of each a trial being safe. The subsequent 30 trials (trials 31-60) took place in a predominantly dangerous environment, with probabilities reversed to.1 safe and .9 threatening. The remaining trials were completed in the safe environment as the initial 30 trials. We maintained the model parameter values consistent with those used in simulation 1 (for the healthy agent) and simulation 3 (for the agent with low self-efficacy).

Simulation results revealed that the agent with low self-efficacy rarely stayed in social settings. The excessive engagement of safety behavior persisted throughout the 100 trials, even when the agent's belief of the safety level had increased (Figure Panel B, right). This discrepancy between cognition and behavior mirrors the findings from Simulation 3. Due to limited social exposure, the low self-efficacy agent failed to update their belief about the social environmental state when the social context became dangerous, which might increase the chance of social mishaps should the agent choose to stay. This observation underscores the importance of appropriate levels of social exposure for achieving accurate belief and adaptive behavioral outcomes. In contrast, the healthy agent demonstrated a more flexible approach, staying in the safe social environments and quitting when the environment became threatening. This adaptive behavioral strategy facilitated more accurate belief updates about the environmental state and resulted in greater social gains. These findings highlight the complex interplay between self-efficacy, perception of the social context, and behavioral adaptation. The simulation results suggest that to increase adaptability, interventions for individuals with low self-efficacy need to address both the cognitive and behavioral aspects.

Supplmentary Figure S1.

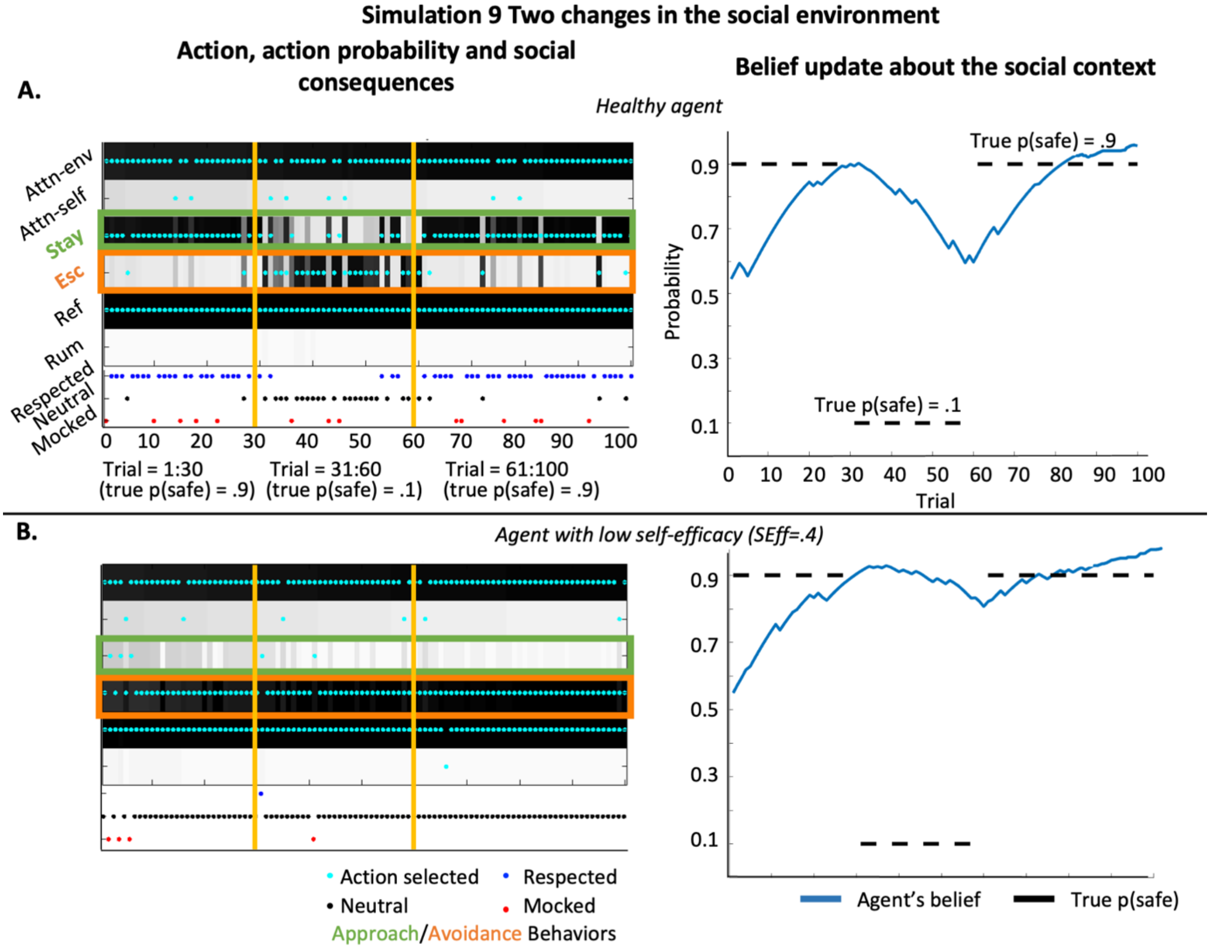
Action selection, perceived social consequences and belief update for agent with altered learning rates.



Note. The selected action, action probability, perceived social consequences and belief update about the social context across the 100 trials for (A) an healthy agent and (B) an agent with low self-efficacy in an social environment that changed twice. The orange lines indicates the point when the social environment changed first from supportive (the probability of the context being safe = .1) to hostile (the probability of the context being safe = .9) (first yellow line at trial 31) and then back to supportive (second yellow line at trial 61). For all panels, darker shades represent higher probability values. Cyan dots mark the actual states for each trial. Blue, black and red dots indicate socially successful, neutral and unsuccessful outcomes observed by the agent, respectively. Attn-env = Attending to the environment, Attn-self =Attending to self, Esc = Escape, Ref = Realistic reflecting, Rum = Rumination. Blue lines depict how the agent's belief about the social context evolved throughout 100 trials. Black lines mark the true probability of the social context being safe.

Supplmentary Figure S2.

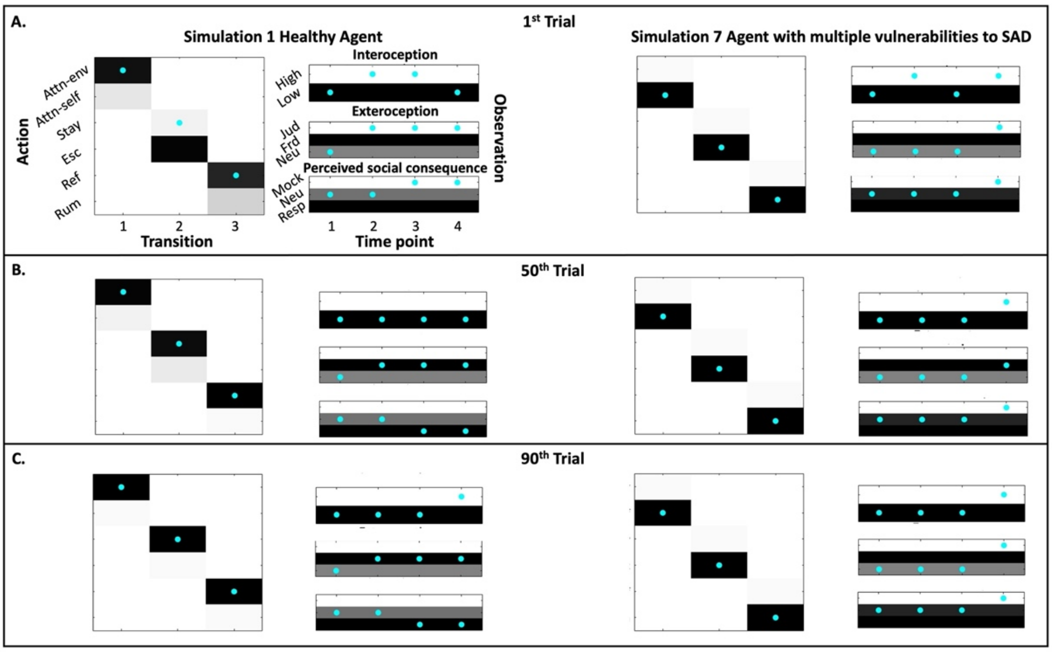
Action selection, perceived social consequences and belief update for agent in a relatively volatile environment.



Note. The selected action, action probability, perceived social consequences and belief update about the social context across the 100 trials for **(A)** an healthy agent and **(B)** an agent with low self=efficacy in an social environment that changed twice**.** The orange lines indicates the point when the social environment changed first from supportive (the probability of the context being safe = .1) to hostile (the probability of the context being safe = .9) (first yellow line at trial 31) and then back to supportive (second yellow line at trial 61). For all panels, darker shades represent higher probability values. Cyan dots mark the actual states for each trial. Blue, black and red dots indicate socially successful, neutral and unsuccessful outcomes observed by the agent, respectively. Attn-env = Attending to the environment, Attn-self =Attending to self, Esc = Escape, Ref = Realistic reflecting, Rum = Rumination. Blue lines depict how the agent's belief about the social context evolved throughout 100 trials. Black lines mark the true probability of the context being safe.

Supplmentary Figure S3.

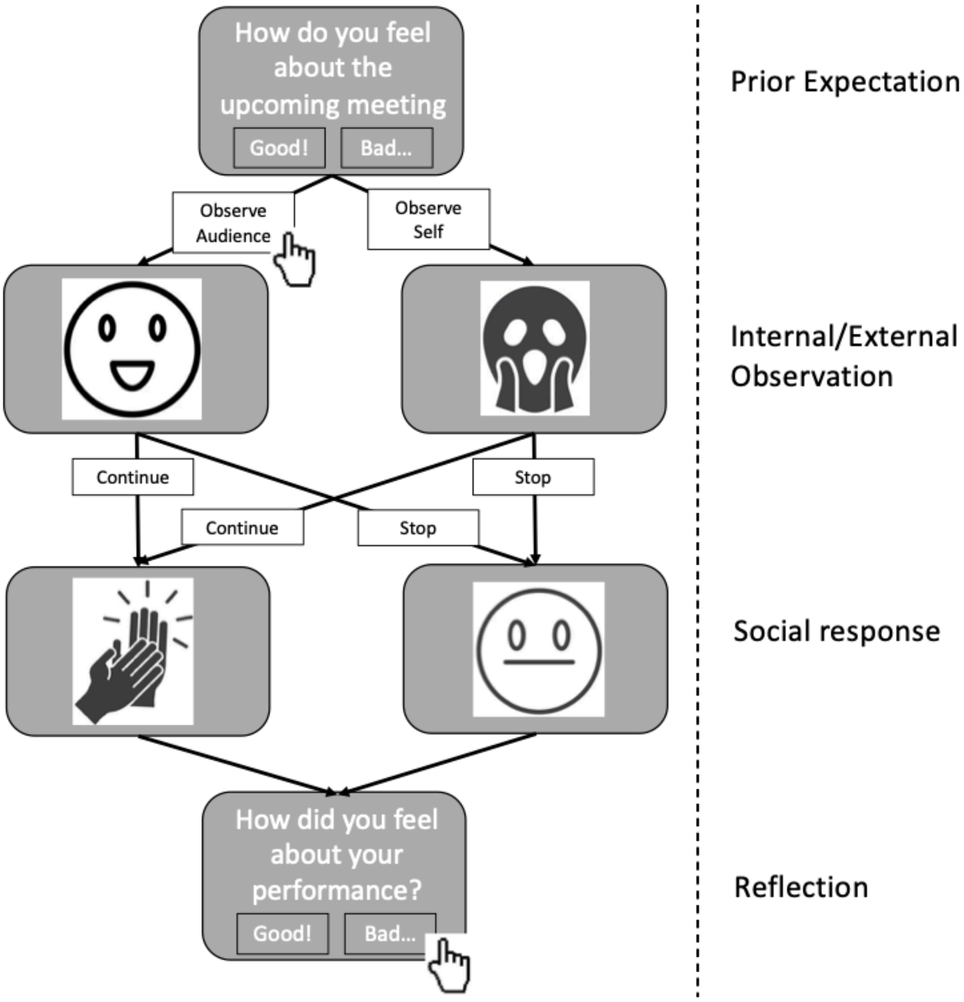
Sampled actions and sensory observations from their probability distributions for simulations 1 and 7.



Notes. For action probabilities (3 panels on the left for each simulation), darker shades represent higher probability values for selecting the corresponding action at each transition. Cyan dots mark the actual selected actions for each trial. For sensory observations (3 panels on the right for each simulation), darker shades represent higher prior expectations (i.e. preferences) for observing the corresponding outcomes at each time point. Cyan dots mark the actual observed outcomes. **(A), (B)** and **(C)** shows the sampled actions and observations at the 1st, 50th and 90th trials respectively.

Supplmentary Figure S4.

A proposed experiment paradigm for model validation and comparison in future research



Note. Sample experimental paradigm for investigating psychopathological pathways for social anxiety disorders using CBT-informed active inference modling. Participants complete multiple trials of giving presentations to different audiences in the experiment. Each trial consists of four phases: (1) reporting prior expectations, (2) presenting while choosing between attending to internal sensations or external audience cues, (3) observing audience feedback, and (4) updating beliefs about performance. Self-report and behavioral choices can be used for model fitting. This design enables quantitative assessment of belief updating processes about social environments and decision-making through the proposed active-inference modelling approach.