Acute isolation is associated with increased reward responsiveness in human adolescents

Livia Tomova

lt503@cam.ac.uk

Department of Psychology, University of Cambridge

Emily Towner

University of Cambridge

Kirsten Thomas

Department of Psychology, University of Cambridge

Sarah-Jayne Blakemore

University of Cambridge https://orcid.org/0000-0002-1690-2805

Article

Keywords:

Posted Date: March 30th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2718114/v1

Additional Declarations: There is **NO** Competing Interest.

Abstract

Social connection is a basic human need and particularly important during adolescence. How a lack of connection impacts adolescent behaviour is unclear. To address this question, we employed experimental short-term isolation, first, to assess how isolation affects reward seeking and reward learning in adolescents aged 16-19 years and, second, whether virtual interactions remediate isolation effects. Isolation was associated with faster decisions to exert effort for rewards and higher reward learning, especially from social feedback. These effects were stronger in participants who reported higher levels of loneliness following isolation. Virtual interactions remediated effects only partially and were associated with lower learning from social feedback. We explored predictors of sensitivity to isolation and found that participants with lower neural reward sensitivity at baseline showed stronger effects of isolation. These results demonstrate that, in adolescents, isolation is associated with higher reward responsiveness, a key driver of motivation and decision-making.

Introduction

Social connection is a basic human need[1,2]. However, we are not always able to obtain the kind of social contact that satisfies our social needs. Reports on loneliness, the subjectively perceived lack of social connection[1], have shown that, across the world, adolescents are the age group most affected by increasing levels of loneliness[3–5]. Over the past two decades, longitudinal research has revealed associations between adolescent loneliness and depression[6–11] and drug use[12–15]. However, the direction of causality is unclear, and bi-directional effects between loneliness and mental ill health have been reported[8]. Experimental designs enable a better understanding of causality, but such approaches have so far been limited to studies in animal models. These studies have consistently shown that isolating adolescent animals primarily affects dopaminergic brain reward circuits and alters reward seeking and reward learning[16,17]. It is unclear to what extent these findings can be translated to human adolescents. As well as reporting high levels of loneliness, young people also report more social media use than other age groups[18]. It has been proposed that there might be a connection between the use of virtual social interactions and loneliness[19], while others have proposed that virtual social interactions might be a potential remedy for rising chronic isolation and loneliness[19]. The displacement hypothesis[20,21] proposes that virtual social interactions displace real-life social interactions, increasing social isolation and loneliness. An alternative, although not mutually exclusive, stimulation hypothesis[22,23] proposes that virtual social interactions offer opportunities to strengthen existing real-life relationships and create new, effective and supportive relationships.

Here, we employed an experimental approach[24] using social isolation of up to four hours in human adolescents to induce subjective feelings of loneliness and assess how they are associated with changes in reward processing (Fig. 1). Participants first underwent a baseline session without any isolation during which we measured reward seeking and reward learning. Subsequently, participants underwent two isolation sessions (order counterbalanced across participants): a total isolation session (iso total), which

involved isolation from any real-life and virtual social interactions, and an isolation session in which participants were isolated from in-person social interactions but had access to any form of virtual social interactions via phone and/or other devices, using any platform they wanted (iso media).

This design allowed us to compare effects of both types of isolation to a baseline unaffected by experience of isolation, and compare effects of being totally isolated to being isolated with access to virtual interactions while keeping other factors (such as spending time alone in a room) constant (see Methods section for details on procedures during isolation). Using this approach, we asked two main questions: (Q1) Is isolation associated with changes in reward seeking and reward learning? and (Q2) Does access to virtual social interactions remediate the effects of isolation on reward processing? If virtual social interactions are sufficient for fulfilling social needs, they should remediate the effects of social isolation similarly to real-life social interactions. Finally, we also sought to explore whether trait neural reward sensitivity predicted individual differences in sensitivity to isolation.

Results

At baseline and after each isolation session, participants underwent a reward seeking (effort-based decision making - EBDM) and a reward learning (probabilistic reinforcement and reversal learning - RL) task. During the baseline session, participants also underwent functional magnetic resonance imaging (fMRI) to provide a measure of neural reward sensitivity which was used in preliminary analyses testing for predictors of individual differences in sensitivity to isolation. All procedures, methods and hypotheses were preregistered on the Open Science Framework (OSF; https://osf.io/kbgsv*).*

All participants (N = 40; age range = 16–19 years; mean age = 17.1; std = 0.9; 22 female) were socially well-connected in that they reported having frequent social interactions (number of face-to-face or virtual interactions that were primarily social in nature in the past month: mean 35.8 (SD 31.9); minimum 10) and several close relationships (mean 7.5 (SD 4.2); minimum 2). Participants reported average levels of pre-existing loneliness for their age group[25,26] (University of California, Los Angeles (UCLA) loneliness scale: mean 35.8 (SD 6.2); maximum 48 out of 80).

Adolescents Show Increased Loneliness And Decreased Positive Mood Following Isolation

Every hour during the two isolation sessions, participants received an electronic reminder to self-report loneliness, mood, boredom and social craving (see methods section *Questionnaires* for details on the selection of timing of ratings). After three hours in both isolation sessions, compared with at the start of the session, participants reported significantly increased loneliness (iso total: $t(39) = 5.87$, $P < 0.001$, Cohen's $d = 1.06$; iso media: $t(39) = 4.59$, $P < 0.001$, $d = 0.68$; Fig. 2) and decreased positive mood (iso total: $t(39) = -6.41$, $P < 0.001$, $d = 0.75$; iso media: $t(39) = -7.05$, $P < 0.001$, $d = 0.82$). The increase in loneliness over time in isolation was higher in the iso total session than in the iso media session (beta (b) = 3.48, t = 2.68, 95% confidence interval (CI) = 0.93, 6.03, P = 0.008). However, the decrease in positive

mood over time in isolation was not significantly different between the iso media and iso total sessions (b = 0.04, t = 0.11, CI=-0.59, 0.66, P = 0.913). Fig. S4 in the supplementary materials depicts effects of isolation on boredom, social craving, negative mood and state anxiety.

Adolescents Show Increased Reward Seeking Following Isolation

To measure reward seeking, we used an effort-based decision making (EBDM) task with three factors: reward (high/low), effort (high/low) and context (social/non-social). Response times (RTs) while participants chose whether to undergo an effortful task have previously been shown to be indicative of strength of preference[27] and are here used as the main measure of reward seeking (see Methods for details). There was a 4-way interaction (b = 0.84, t = 2.45, CI = 0.17, 1.51, P = 0.015) between session (baseline, iso total, iso media), reward (high, low), context (social, non-social) and effort (high, low) on RTs in the EBDM task (the time it took participants to decide whether to complete a trial). Pairwise comparisons (Bonferroni corrected $p < 0.0063$ (0.05/8)) showed that participants were significantly faster in deciding to undergo trials in the iso total session compared with baseline, when rewards were high (all p-values < = 0.005, range Cohen's d = 0.60–0.82); see Fig. 3 and Fig. S1 and Table S3 in the supplementary materials for full results). Mean RTs for low reward conditions were only faster in the iso total compared to baseline session when effort was low and only in the social context (t(28)=-3.48, P = 0.002, d = 0.71; all other p-values were above the Bonferroni corrected alpha level of 0.01). This indicates that, after total isolation, participants were specifically faster in trials where rewards were high instead of showing generally decreased response times across all conditions.

Note that the number of trials from which RTs were calculated was lower in the low reward conditions because participants decided to carry out fewer low reward trials. None of the RTs in the iso media session were significantly different from baseline (all ps > 0.008).

We performed sensitivity analyses to assess whether order effects might have contributed to changes in response times between sessions (see supplementary materials *Sensitivity Analyses Reward Seeking*) and found no evidence that order effects were driving the main result of decreased response times to high rewards after total isolation.

Next, we explored whether self-reported loneliness at baseline and after three hours of isolation (in the iso total and iso media session) interacted with session effects by adding it as a predictor in a model with the three factors described above. We found a main effect of loneliness (b=-0.01, t=-2.70, CI=-0.02, -0.004, P = 0.007), indicating that, across all sessions, higher feelings of loneliness were associated with faster response times in the EBDM task (Fig. 4). We also found an interaction between context and loneliness in the iso media session (b = 0.27, t = 2.08, CI = 0.01, 0.53, P = 0.039), indicating that participants who reported higher loneliness showed slower response times in the social context in the iso media session (Fig. 4). We also tested whether self-reported boredom at baseline and after three hours of isolation (in the iso total and iso media session) interacted with session effects to assess whether effects of session might have been driven by unspecific effects of boredom during isolation. Self-reported boredom did not

show any main effects or interactions with effects of session on RTs in the EBDM task (see supplementary materials *Exploratory analyses* for results).

Adolescents Show Increased Reward Learning Following Isolation

We used a probabilistic reinforcement and reversal learning task[28] to measure how isolation affects the ability to learn stimulus–reinforcement associations, and to reverse them, based on probabilistic feedback. Feedback was given either via social cues (facial expressions) or non-social cues (symbols). Data analysis involved, first, selecting a model that best captured participants' behaviour by comparing model fits between three different reversal learning models for each session (baseline, iso total, iso media) and condition (social, non-social). Three models were compared: (i) a reward-punishment model[29], which is based on the classic Rescorla-Wagner model of conditioning[30], but expanded with separate learning rates for reward and non-reward trials, here treating non-wins as non-rewards; (ii) an experience-weighted attraction model[31], which is also based on the Rescorla-Wagner model, but includes a stickiness parameter that captures perseveration bias (which can artificially drive learning rate asymmetries[32]). This model also includes an 'experience' weight parameter that decouples acquisition and reversal; and (iii) a fictitious update model[33], which is an approximation to the Hidden Markov Model formulation based on Hampton et al. 2006[34] but which has been further extended[33,35] to include an update rule for the unchosen option, taking into account the knowledge individuals gain about the unchosen option, here with separate learning rates for positive prediction errors (learning_pos; outcome better than expected) and negative prediction errors (learning_neg; outcome worse than expected). See Methods for a full description of all models and model fitting.

We found that model fit was best for the fictitious update model for each session and condition (see Table S5 in the supplementary materials for results of model fitting). We then assessed the model fit of the winning model (i.e., the fictitious update model) by running posterior predictive checks. To do this, we generated simulated data for each participant using the fitted model and checked if the predictions were similar to the actual data from each participant. Overall, the model predictions captured the real data well (see Fig. S3 in the supplementary materials for examples).

Population-level Parameters

As a next step, we extracted the population-level mean posterior distributions (i.e., mean across all participants' posterior distributions) of each parameter estimated by the model: learning rate (learning_pos and learning_neg) and inverse temperature (beta; a measure of exploration). To compare the posterior distributions between the different conditions, we then calculated the pairwise difference between the mean posterior distributions for each parameter and each condition. Figure 6 depicts the mean posterior distributions for each parameter (Fig. 6A) and the difference distributions between each parameter in each session for each feedback type (Fig. 6B-D). For the difference distributions between the

different sessions, we calculated the posterior probability of μ1 − μ2 > 0 (sum values > 0 / total number of values) in which 0 indicates no difference between conditions. A probability of 50% for a difference in either direction indicates chance.

Learning_pos rates (positive prediction errors - outcome better than expected). There was a 79% probability for a *higher* learning rate from social feedback, and a 57% probability for a *higher* learning rate from non-social feedback, in the iso total compared to baseline session. There was a 61% probability for a *lower* learning rate from social feedback, and a 52% probability for a *lower* learning rate from non-social feedback, in the iso media compared to baseline session. For the difference distribution between iso total and iso media, there was a 90% probability for a *higher* learning rate from social feedback in the iso total compared to the iso media session and a 60% probability for a *higher* learning rate from non-social feedback. Thus, in sum, the results indicate a higher probability for higher learning_pos rates from social feedback in the iso total compared to baseline and iso media session.

Learning_neg rates (negative prediction errors - outcome worse than expected). There was an 86% probability for a *higher* learning rate from social feedback, and a 62% probability for a *higher* learning rate from non-social feedback, in the iso total compared to baseline session. There was a 74% probability for a *higher* learning rate from social feedback, and a 67% probability for a *higher* learning rate from non-social feedback, in the iso media compared to baseline session. For the difference distribution between iso total and iso media, there was a 65% probability for a *higher* learning rate from social feedback in the iso total compared to the iso media session and a 57% probability for a *lower* learning rate from non-social feedback. In sum, the results indicate a higher probability for higher learning_neg rates from social feedback in both isolation sessions (iso total and iso media) compared to baseline.

Beta (inverse temperature - a measure of exploration). There was a 66% probability for a *higher* beta in the social feedback condition, and an 86% probability for a *higher* beta in the non-social feedback condition, in the iso total compared to baseline session. There was a 99% probability for a *higher* beta in the social feedback condition, and a 94% probability for a *higher* beta in the non-social feedback condition, in the iso media compared to baseline session. For the difference distribution between iso total and iso media, there was a 99% probability for a *lower* beta in the social feedback condition in the iso total compared to the iso media session and a 72% probability for a *lower* beta in the non-social feedback condition. In sum, the results indicate a higher probability for higher beta (indicating less exploration) in the iso media session compared to baseline and the iso total session, regardless of feedback type.

Individual-level Parameters

Finally, following methods of Crawley et al.[36], we extracted learning rate (learning_pos and learning_neg) and inverse temperature (beta; a measure of exploration) for each participant to assess effects of session and condition on these parameters. There was a significant session-by-condition interaction for the learning_pos parameter (b = 0.20, t = 31.93, CI = 0.18, 0.21, P < 0.001), indicating that participants showed higher learning rates in the iso total session compared with baseline, and that this effect was stronger for

learning from social feedback compared with non-social feedback. The session-by-condition interaction for the iso media session (compared with baseline; b=-0.04, t=-7.12, CI=-0.06, -0.03, P < 0.001) showed a different pattern: there was a *decrease* in learning from social feedback following the iso media session compared with baseline, and no effect on learning from non-social feedback.

Pairwise comparisons (Bonferroni corrected p < 0.0083 (0.05/6)) showed that learning_pos rates were significantly different for each feedback type in the iso total compared with baseline (social: t(39) = 56.72, P < 0.001, d = 12.38; non-social: t(39) = 5.10, P < 0.001, d = 1.11) and compared with the iso media session (social: t(39) = 77.42 P < 0.001, d = 17.28; non-social: t(39) = 8.22, P < 0.001, d = 1.83). Learning_pos rates did not differ between the iso media and baseline session for non-social feedback (t(39)=-0.82, P = 0.415, d = 0.18), while learning_pos rates decreased in the iso media session compared with baseline for social feedback (t(39)=-10.72, P < 0.001, d = 2.36).

Learning from negative prediction errors (learning_neg) showed an effect of condition, in that participants overall learned better from negative social feedback compared with non-social feedback (b = 0.06, t = 2.25, CI = 0.01, 0.11, P = 0.025), but there was no effect of session.

Beta was overall higher (indicating less exploration) in the iso media compared to baseline session (b = 0.26, t = 2.62, CI = 0.07,0.46, P = 0.009) but showed no other effects or interactions with session. However, sensitivity analyses showed that, when controlling for session order and days between sessions, beta was also significantly lower in the iso total compared to baseline session (see supplementary materials for details).

To interrogate what these changes in parameters mean for how well participants performed in the task, we used simulations to assess the optimal learning rate and inverse temperature for our specific task environment (see Methods for details). As expected[37] for a short task (28 trials per block) with frequent changes of reward contingencies (every 7 trials), a high learning rate (indicating stronger reliance on more recent outcomes for value updating) was optimal. Beta (inverse temperature) did not strongly affect performance at values above 2.5. Lower values (indicating high exploration; Fig. S2) were associated with poor performance. Thus, for the current task, frequently updating reward contingencies was associated with better performance.

We performed sensitivity analyses to assess whether order effects contributed to changes in parameters over sessions (see supplementary materials *Sensitivity Analyses Reward Learning*) and did not find evidence that order effects were driving the main result of higher learning_pos rates from social feedback after total isolation.

Next, we explored whether self-reported loneliness after three hours of isolation interacted with effects of session on learning_pos rates by adding it as a predictor in a model with session and condition. We found an interaction between session, condition and loneliness for both isolation sessions (iso total: b = 0.001, t = 7.53, CI = 0.0006, 0.0011, P < 0.001; iso media: b = 0.001, t = 7.73, CI = 0.0007, 0.0012, P < 0.001), indicating that higher feelings of loneliness were associated with higher learning_pos rates in response to

social cues in the RL task (Fig. 7). Again, we also tested whether self-reported boredom at baseline and after three hours of isolation (in the iso total and iso media session) interacted with session effects to assess whether effects of session might have been driven by unspecific effects of boredom during isolation. Self-reported boredom did not show any main effects or interactions with effects of session on learning_pos rates (see supplementary materials *Exploratory analyses* for results).

Exploratory analyses of gender differences showed an effect of gender on inverse temperature in the reward learning task indicating higher exploration in males compared to females (b=-0.29, t=-3.21, CI=-0.46, -0.11, P = 0.002) but no interactions between gender and isolation (see supplementary materials for full results). We also explored whether pre-existing trait anxiety, depression scores, chronic loneliness, as well as boredom during the sessions, interacted with the effects of isolation. We found that participants who reported higher chronic loneliness showed stronger effects of social isolation on RTs in the social context trials in the EBDM task (b=-0.07, t=-2.58, CI=-0.12, -0.02, P = 0.011; details in supplementary materials).

Preliminary Evidence That Neural Reward Sensitivity At Baseline Predicts Effect Of Isolation On Reward Seeking

Each participant underwent a functional magnetic resonance imaging (fMRI) scan at baseline, during which they carried out four runs of a monetary incentive delay (MID) task. This task is commonly used to study activity in the neural reward system[38]. Figure 8a shows the group level results from the task (contrast high reward anticipation > no reward anticipation), which revealed activity in reward-related regions, as expected[38].

Our preregistered analyses comprised two measures of individual 'neural reward sensitivity': a univariate measure for which we extracted the mean activity for the contrast high reward anticipation > no reward anticipation for each participant; and a multivariate measure, which was the prediction accuracy of a classifier trained to classify neural patterns of high rewards from no rewards during the anticipation phase (see Methods for details).

We assessed both the univariate and multivariate measures of neural reward sensitivity in *a priori* selected regions of interest (ROIs), which have been consistently activated during anticipation of rewards[38]. These regions included: the ventral striatum (VS), midbrain (MB), amygdala (AMY), anterior insula (AI), occipital cortex (OC), thalamus (TH) and supplementary motor area (SMA). We then tested whether neural reward sensitivity correlated with reward seeking (RTs from the high reward / high effort conditions in the EBDM task; mean across social and non-social contexts) and reward learning (learning_pos rates in the RL task; separate analyses for learning_pos rates from social and non-social feedback) in the isolation session. We calculated separate correlations for each ROI for each measure of neural reward sensitivity (univariate and multivariate; 14 tests in total) and report results as significant at $p < 0.003$ (0.05/14).

Univariate neural reward sensitivity in the OC showed a significant correlation with RTs in the EBDM task in the iso total session ($r(38) = 0.53$, $p < 0.001$; Fig. 8b). This association was not present at baseline ($r(38)=-0.096$, $p = 0.57$; correlations significantly differed from each other: $z = 2.16$, $p = 0.031$), suggesting that neural reward sensitivity specifically predicted reward seeking following isolation rather than general reward seeking. Multivariate neural reward sensitivity did not predict reward seeking in the iso total session for any of the ROIs (all p-values > 0.30).

Neural reward sensitivity (univariate or multivariate) did not predict reward learning in either the iso total session or the iso media session for any of the ROIs for either social or non-social feedback (all p-values > 0.03).

Discussion

In adolescents aged 16–19 years, short-term social isolation was associated with an increase in reward responsiveness measured in two different tasks. First, we found that, even after a relatively short period of isolation (four hours or less), participants reported feeling lonelier at the end of the isolation period than they did at the beginning; this was true for both total isolation and isolation in which participants were able to communicate with others virtually. Second, the results from the effort-based decision making (EBDM) task showed that total isolation (but not isolation with virtual social interactions) was associated with faster response times when participants were deciding whether or not to undergo an effortful task for high rewards, which is indicative of increased reward seeking[39]. Third, the results from the reward learning (RL) task revealed that total isolation (but not isolation with virtual social interactions) was associated with an increase in learning from rewards, particularly when learning from social feedback. Our findings are partially consistent with results from animal models, which have shown that social isolation in adolescence increases reward responsiveness[17]. In animals, social isolation increases responsiveness to social rewards[40] seeking of food or drug rewards and risk of developing addictions[41]. Our findings suggest that there is a similar effect of increased responsiveness to rewards – both social and non-social – immediately following total isolation in human adolescents. Furthermore, subjective feelings of loneliness during each sessions predicted higher reward seeking overall (regardless of session), and higher learning rates to social feedback. Self-reported boredom, on the other hand, did not predict any effects of our tasks (see supplementary materials for results). Thus, these findings suggest that isolation-induced changes in reward seeking and reward learning were driven by experienced loneliness rather than unspecific effects of boredom during isolation.

The finding that total isolation was associated with higher reward learning is contrary to results from animal studies. Animal studies have found that isolation leads to perseveration to initially learned reward-outcome associations following reversals[42,43]. In contrast, in our study, adolescents constantly updated associations between cues and rewards throughout the task, particularly in response to social feedback. Importantly, however, while a high learning rate was optimal in our task, high learning rates are not always optimal and the interpretation of a learning rate depends on the specific task environment[37]. A

high learning rate indicates that participants only use the most recent outcomes for value updating, while a low learning rate indicates that both recent outcomes and previous outcomes contribute to the value update[37]. Thus, a low learning rate can be optimal in task environments in which reward contingencies are more stable. Considering this, our results suggest that, following total isolation, when feeling lonelier, adolescents place more weight on each instance of social feedback they receive to guide their behaviour, compared with at baseline, when they feel more socially connected. This could be beneficial in cases when social feedback is informative but could also lead to sub-optimal behaviour when social feedback is less informative.

Participants reported feeling lonelier at the end of the total isolation period than they did at the beginning and, interestingly, this was also true when adolescents had access to virtual social interactions during isolation. However, the increase in self-reported loneliness was lower than following total isolation. In contrast, positive mood decreased to a similar extent in both isolation sessions. Whether virtual social interactions fulfil social needs is debated[19] and some findings suggest the contrary, that is, that virtual social interactions can increase loneliness[19]. The current study suggests that virtual social interactions can partially – but not completely – remediate feelings of loneliness associated with isolation.

In line with this, in the iso media session, adolescents showed no significant difference in reward seeking (in the EBDM task) compared with baseline, nor did they show increased reward learning (in the RL task), as was seen in the iso total session. These findings suggest that virtual social interactions mitigate the effects of isolation on reward processing. However, contrary to our predictions, having access to virtual social interactions during isolation led to decreased reward learning from social feedback in the RL task. While access to virtual social interactions might have general effects on attention or motivation, the fact that learning was only affected on trials when feedback was given via social cues suggests a possible alternative explanation. A substantial number of participants (18 out of 40; 45%) reported spending more than 50% of their isolation time participating in virtual social interactions. Recent research has shown that human behaviour on social media platforms conforms to the principles of reward learning: users spaced their posts to maximize the average rate of accrued social rewards, considering both the effort cost of posting and the opportunity cost of inaction[44]. Thus, a period of intense engagement in virtual social interactions might specifically decrease the reinforcing nature of other social cues, resulting in lower learning from such cues. Therefore, the decreased social learning in the iso media session might have been caused by intense engagement in virtual social interactions. However, these interpretations are speculative, and the findings require replication and further specification.

We found preliminary evidence that neural reward sensitivity in the occipital cortex measured at baseline predicted individual differences in the effects of isolation on reward seeking: adolescents with lower neural reward sensitivity in the occipital cortex at baseline showed increased reward seeking in the EBDM task after total isolation compared to baseline. Conceptually, low responsiveness to rewards has been linked to anhedonia and might represent a marker for stronger sensitivity to stressors[45]. Previous research has shown that the magnitude of neural activation during anticipation of rewards can predict

real-world responsivity to stressors[46,47]. For example, adolescents who showed lower neural reward sensitivity in the MID task at age 10 reported higher stress levels in response to a laboratory stressor at age 13[46]. Similar associations between neural reward sensitivity and stress reactivity have been shown in adults[47]. Our findings suggest that lower neural responsivity to rewards might be indicative of higher sensitivity to the stress of isolation. While the occipital cortex has been identified in meta-analyses as one of the core brain regions consistently activated during reward anticipation in the MID task[38] and in other monetary reward tasks[48], it is unclear what exact neural mechanism underlies this fMRI signal. Given that the high reward and no reward condition did not differ visually in the MID task, except for the number presented inside the star ("0" in the no reward condition and "5" in the high reward condition), the measured activity might represent a signal of visual attention[49] towards cues predicting rewards. Thus, individual differences in such attention signals to rewards might represent the neural mechanism underlying the contribution of the occipital cortex to the behavioural change.

Our exploratory analyses showed gender differences in the reward learning task indicating higher exploration in males compared to females but no interactions between gender and isolation. Previous animal research has shown higher exploration during learning in males[50] and our findings indicate a similar effect in humans.

The present study raises key questions for future research. First, the neural processes underlying isolation-induced changes during reward seeking and reward learning remain to be explored. Our study suggests that, following isolation, adolescents show increased behavioural responsiveness to rewards, driving them to seek out rewards more and learn reward contingencies better, particularly from *positive* feedback, and particularly when that feedback is *social*. We do not know, however, whether these effects are accompanied by increased responsiveness of the brain's reward system during these tasks. While increased activity in reward circuitry has been consistently shown in animal model studies of adolescent isolation, direct experimental assessment in humans would provide further evidence of isolation-induced increases in reward responsiveness.

Second, whether our experimental findings can be translated to the effects of real-life chronic loneliness is unknown. Although the experimental isolation successfully induced feelings of loneliness, the manipulation was short, and adolescents knew they could re-engage in social interactions afterwards. Importantly, the relatively short duration of isolation employed in the current study would not be expected to induce any longer-term effects on adolescent behaviour.

Chronic and acute loneliness can have opposite effects on social behaviour[51]. The social homeostasis model suggests that prolonged engagement of mechanisms intended for short-term adaptation to acute loneliness, such as hypervigilance and increased social motivation, can cause pathological states during chronic loneliness[52]. According to this view, a prolongation of increased reward responsiveness following acute loneliness during chronic loneliness might ultimately result in imbalances of the reward system, increasing proneness to developing compulsive behaviours, or a blunting of reward system activity.

Indeed, longitudinal studies in humans consistently link loneliness to depression and substance use[6–15]. Here, there was an interaction between chronic and acute loneliness on reward seeking: participants who reported higher pre-existing chronic loneliness showed stronger effects of isolation on reward seeking when rewards were presented in social contexts. It should be noted that chronic loneliness was relatively low amongst the participants in the current study as young people with high pre-existing loneliness were not eligible to take part. Further assessment of the relationship between acute and chronic loneliness should address fundamental questions concerning the nature, origins and interrelations between loneliness and mental health problems in adolescence.

During adolescence, brain reward circuits undergo critical remodeling[53]. The domain of reward processing is of particular importance due to its powerful effects on behaviour[54] and mental health[55]. The present study suggests that brief durations of social isolation in adolescence alter core components[54] of reward processing (i.e., reward seeking and reward learning) in human adolescents. Our findings indicate that real-life isolation (such as during a pandemic or when used as punishment in schools) might result in two key changes in adolescent behaviour. First, real-life isolation might result in higher seeking of rewards, such as food or recreational drugs. Second, the stronger reliance on social feedback during learning following isolation suggests that real-life isolation might lead to increased social influence in adolescence.

Methods

Participants

We collected data from 42 participants; 2 participants were unable to complete all experimental sessions and so were dropped from analysis, leaving 40 complete datasets (mean age = 17.1; std = 0.9; 22 female). Participants were recruited through online advertisements and flyers. The study description indicated that we were interested in studying how different forms of spending time alone (with or without access to virtual interactions) affects adolescent cognition. Interested individuals completed a screening questionnaire to assess eligibility for the study (questionnaire data was collected using Qualtrics and REDCap[56] Software). Participants were asked to indicate their gender (how they describe themselves) and sex (assigned at birth). In our sample of participants, gender labels were not different from sex labels and we use the term gender throughout the main text and supplement.

Participants were eligible to take part in the study if they were between 16–19 years of age, and reported no permanently implanted metal in their body, no history of brain damage and no currently diagnosed mental health disorder. Because data for this study was collected during a COVID-19 pandemic, we also followed exclusion criteria based on a departmental COVID-19 risk assessment. These criteria excluded participants with chronic underlying health conditions (such as asthma), participants who currently felt ill (or had tested positive for COVID-19), and participants who smoked. As we aimed to study effects of isolation in a sample of adolescents who have frequent and regular social interactions, we also excluded

people who: (1) lived alone; (2) reported high feelings of chronic loneliness on the UCLA loneliness scale[26] (we excluded adolescents with scores > 50, which is 2 standard deviations above the mean for an adolescent sample[57]); and/or (3) reported substantially smaller social network sizes than previously reported for an adolescent sample[58] measured via two measures: (a) Number of close friends (the original questionnaire asks for *number of people who give social support*[59], which we adapted to *number of close friends* to simplify the question for our adolescent sample); and (b) Number of social interactions in the past month: counting face-to-face and virtual social interactions that were primarily social in nature (i.e., excluding professional interactions, like talking to a doctor, teacher, or hairdresser). We excluded participants who reported fewer than two close friends and fewer than 10 social interactions in one month (which is ~ 7 standard deviations below the previously reported means for adolescents for both measures[58]). The exclusion thresholds for the social network size measures were lower than previously reported[24] because data for this study was collected during the COVID-19 pandemic, which required social distancing and therefore lower levels of social connectedness were expected throughout the population.

All experimental procedures were approved by the Psychology research ethics committee at University of Cambridge, UK. Participants signed a consent form describing all experimental procedures before participating in the study. Each participant was compensated with up to £127 (minimum payment for each participant was £107) for participating in three sessions.

Experimental procedures

Each participant underwent three experimental sessions, separated by at least 24 hours. Figure 1 shows an overview of the experimental procedures. Each participant first completed a baseline session which included MRI scanning and the behavioural tasks (see *Behavioural: Reward tasks* for details). Following the baseline session, participants were invited to two isolation sessions (order counterbalanced). One session included up to 4 hours of total social isolation (iso total) during which participants had no access to any social interactions (real-life or virtual); another session included up to 4 hours of social isolation with access to virtual social interactions (iso media).

Following the methods of Tomova et al. 2020[24], we sought to experimentally induce the subjective experience of loneliness in adolescent participants. Re-analysing loneliness ratings from a subset of 18–24-year-old participants from Tomova et al. 2020 and pilot measures in two 16-year-olds (N_{total} = 21) showed that two hours of isolation resulted in increased self-reported loneliness in this age group. We therefore reduced the duration of isolation in this study compared with Tomova et al. 2020. The minimal duration of each isolation session was set to 3h 30min and the maximal duration was 4 hours. We randomly assigned an added duration of 0–30 mins of isolation time to each session (in steps of 5 mins, separately for each session), so that participants were not able to predict the precise end of the isolation period in either session. The average isolation duration was similar between sessions (iso total: m = 3h 46 mins, std = 11.2 mins; iso media: m = 3h 47 mins, std = 10.0; t(39)=-0.28, p = 0.79). All participants

underwent all three experimental conditions (baseline, iso total and iso media). Each participant was pseudo-randomly assigned to one order, with the restriction that baseline was always first and that each order was approximately equally likely in the full sample. This design allowed us to compare effects of both types of isolation to a baseline unaffected by experience of isolation, and to compare effects of being totally isolated to being isolated with access to virtual interactions, while keeping other factors (such as spending time physically alone in a room) constant.

Note that in this study power calculations were made for within-subject effects of isolation (see also preregistration) and between-subject analyses (i.e., analyses on the predictors of individual differences in sensitivity to isolation) should be taken as preliminary.

During the iso total session, participants had access to a variety of non-social activities (such as games, puzzles, sudoku, books, drawing/writing supplies). We also encouraged participants to bring activities that they would like to occupy themselves with during the isolation. Examples of what participants chose to bring with them included school homework, a jewellery making kit, art materials and nail polish.

As we aimed to keep all social interactions during this session to a minimum, participants were given extensive instructions about the isolation procedures and the subsequent behavioural paradigms before starting social isolation. Participants gave their phones and laptops to the experimenter and were guided to a room containing an armchair, a desk and an office chair, and a fridge with a selection of food and beverages. The room had a roof window, which allowed day light to enter the room but kept participants from seeing other people outside the building. Participants remained in that room for the duration of the isolation and the subsequent behavioural testing. Participants were told that they could spend the isolation time however they wanted with the restriction that they should fill out the questionnaire every hour and should avoid sleeping during isolation. Participants were reminded to fill out the questionnaires via an alarm sound that went off every hour during isolation and at the end of isolation, indicating to participants that they should start the behavioural tasks. Participants were provided with a desk computer (with parental controls enabled), allowing them to visit only our Slack channel (an online messenger software allowing communication between a group of people (www.slack.com)) and the webpage containing our online questionnaires. Messaging in Slack was restricted to emergencies (that is, in case participants ran into problems that required assistance from the research team during isolation). During isolation and the subsequent behavioural testing, we monitored participants via a live camera, which allowed a researcher to check in on the participant without interacting with them. The camera only provided a live stream of the participant in the room to the experimenter and did not store any recordings of participants. Participants were informed about the camera in an information sheet they received before agreeing to participate in the experiment. Participants completed an online questionnaire (see *Questionnaires* for details) every hour during the social isolation period. The behavioural testing was started immediately after the social isolation. Participants first filled out questionnaires and then started the behavioural tasks, which they started themselves (i.e., there was no interaction with the experimenter between the end of isolation and the beginning of the behavioural testing). After the session, a member

of the research team chatted with the participants about their experiences during isolation and made sure participants were not feeling troubled.

Procedures during the iso media session were the same as in the iso total session, except that participants could bring any electronic device with them into the room and were told they could use them as much as they wanted. Most participants brought their tablet or laptop as well as their phones. See section Questionnaires for descriptive data on how participants reported to have engaged with others virtually during the iso media session.

MRI scanning

MRI data were collected on an ultra-high field 7 Tesla MRI (Siemens 7T Terra) located at the Wolfson Brain Imaging Centre, University of Cambridge. For each participant, an MP2RAGE sequence was used to collect T1-weighted structural images in 224 interleaved sagittal slices with 0.7-mm isotropic voxels (FOV: 224 mm). We also collected a field map (phase-difference B0 estimation; echo time 1 (TE1) = 1.11 ms, echo time 2 (TE2) = 3.06 ms) to control for spatial distortions, which are particularly problematic with ultra-high field scanning[60]. Furthermore, DTI images were acquired with a scanning protocol of 90 1.4 mm-thick contiguous axial slices with an in-plane resolution of 1.4 × 1.4 mm providing full brain coverage. DTI data are outside the scope of this manuscript and will be described in a separate paper.

During acquisition of the anatomical images (~ 7 min in total), participants underwent a practice run of the MID task (see “MID task” for details).

Subsequently, we collected functional data during four runs of the MID task. Each functional run consisted of 250 volumes with 96 T2*- weighted echo planar slices (TR = 1500 ms, TE = 25 ms, FOV = 192 mm, 128 x 128 matrix, yielding a voxel size of 1.5 x 1.5 x 1.5 mm^3) acquired as a partial-head volume in an anteroposterior phase-encoding direction using interleaved slices. The MID task took approximately 25 mins in total.

Functional MRI: Neural Reward Sensitivity

Using functional magnetic resonance imaging (fMRI), we measured blood-oxygen-level-dependent (BOLD) signal at a single voxel in successive scans (a voxel time-series) in response to anticipating rewards of different magnitudes during a monetary incentive delay (MID[61]) task. The MID task is a well-established task used to study neural correlates of reward anticipation (i.e., winning money) and was used here as a marker of neural reward sensitivity. During the task, participants were able to win either £0.2 (small win) or £5 (large win), or lose £−0.2 (small loss), or £−5 (large loss). On some trials, participants did not win or lose any money, which served as control trials. On each trial, participants were first presented with a cue indicating whether they could win or lose money on that trial. The cue was a yellow star with the amount of the possible win (or loss) written inside (i.e., 0 (control trials), 0.2 (small win trials), 5 (large win trials), − 0.2 (small loss trials), − 5 (large loss trials)). Following the cue,

participants saw a white circle on the screen presented for a jittered time interval (range 1.5–4 s, mean 2.5 s) briefly followed by a white square, which was presented for 100 ms and then the circle again. Participants' task was to press a button as fast as possible when they saw the square. Following their response, participants were presented with feedback revealing whether they won (or avoided a loss) or not on that trial (presented for 500 ms). At the end of each trial participants saw a fixation cross for a jittered time interval (range 2–6 s, mean 3.5 s) before the next trial began. The task was adaptive and the time window for how fast participants needed to respond narrowed or widened depending on their response times during the task to ensure that participants won on most (~ 80%) of the trials (to keep the task rewarding) but also that the task was not too easy (to keep participants engaged). Participants underwent a practice run of the task during the anatomical MP2RAGE scan, and the data from this run was used to calibrate the initial response window for the MID task.

Using functional magnetic resonance imaging (fMRI), we measured blood-oxygen-level-dependent (BOLD) signal at a single voxel in successive scans (a voxel time-series) in response to anticipating rewards of different magnitudes during a monetary incentive delay (MID[61]) task. The MID task is a well-established task used to study neural correlates of reward anticipation (i.e., winning money) and was used here as a marker of neural reward sensitivity. During the task, participants were able to win either £0.2 (small win) or £5 (large win), or lose £−0.2 (small loss), or £−5 (large loss). On some trials, participants did not win or lose any money, which served as control trials. On each trial, participants were first presented with a cue indicating whether they could win or lose money on that trial. The cue was a yellow star with the amount of the possible win (or loss) written inside (i.e., 0 (control trials), 0.2 (small win trials), 5 (large win trials), − 0.2 (small loss trials), − 5 (large loss trials)). Following the cue, participants saw a white circle on the screen presented for a jittered time interval (range 1.5–4 s, mean 2.5 s) briefly followed by a white square, which was presented for 100 ms and then the circle again. Participants' task was to press a button as fast as possible when they saw the square. Following their response, participants were presented with feedback revealing whether they won (or avoided a loss) or not on that trial (presented for 500 ms). At the end of each trial participants saw a fixation cross for a jittered time interval (range 2–6 s, mean 3.5 s) before the next trial began. The task was adaptive and the time window for how fast participants needed to respond narrowed or widened depending on their response times during the task to ensure that participants won on most (~ 80%) of the trials (to keep the task rewarding) but also that the task was not too easy (to keep participants engaged). Participants underwent a practice run of the task during the anatomical MP2RAGE scan, and the data from this run was used to calibrate the initial response window for the MID task.

Questionnaires

Baseline. At the beginning of the behavioural testing in the baseline session, participants completed self-report questionnaires assessing trait and state anxiety (the State Trait Anxiety Index (STAI[62])) and depression (the Center for Epidemiological Studies Depression (CES-D) scale)[63]. We report exploratory analyses assessing interactions between individual differences in trait anxiety and depression, in addition

to measures of chronic loneliness (UCLA loneliness scale[26]) collected during the online screening, and the effects of our experimental manipulations in the supplementary text. Measures of self-reported loneliness (0-100), social craving (0-100), boredom (0-100) and current affect (using the Positive and Negative Affect Schedule (PANAS[64]) were also acquired at the beginning of the baseline session. In addition, participants were asked whether they drink alcohol and whether they vape (at least once per month). If yes, participants were asked to rate their current urge to drink alcohol and to vape on a scale from (0) not at all to (100) extremely. In accordance with the University COVID-19 risk assessment, participants who reported that they smoke were excluded from participation, and therefore we did not ask about craving in the domain of smoking. Because only 55% of our participants (N = 22) reported that they drank alcohol at all, and only two participants reported that they vaped, there was not sufficient power to assess effects of isolation on alcohol or vape craving.

Isolation sessions. Following each isolation session (iso media and iso total), immediately before participants started the behavioural tasks, we assessed state anxiety and craving for alcohol and vaping. In addition, participants self-reported loneliness, social craving, boredom and current affect at the beginning of each isolation session (T0) and then after every hour of isolation for 3 hours (T1, T2, T3), followed by a final rating at the end of isolation (T4; 3 hours 30 min to 4 hours of isolation, varying between participants). Effects of isolation on state anxiety are reported in the supplementary text.

All statistical analyses using self-report questionnaires were conducted on measures taken after 3 hours (T3) of isolation. This allowed us to capture participants' affective state after a substantial period of isolation but before they knew it was over. Plots including loneliness and mood ratings at the final time point (T4) are shown in Fig. S5, and statistics using these time points are included in the supplementary text. Note that the general results remain the same regardless of whether T3 or T4 ratings were used. Boredom measures were used in exploratory analyses to assess whether self-reported boredom interacted with the effects of isolation (see supplementary text).

After the iso media session, participants completed a questionnaire which asked them to report how much they engaged in virtual social interactions during the session (percentage of time spent engaging in virtual social interactions during session (0-100%)). We also asked participants to indicate which method(s) (texting/messaging, voice calls, video calls, commenting/posting, gaming, other) and platform(s) (Instagram, Facebook, Facebook messenger, Snapchat, TikTok, Twitter, WhatsApp, other) they mostly used for their virtual social interactions during the session. Participants were asked to list whom they interacted with virtually (i.e., friends, family, acquaintances, romantic partner, other). Participants could select multiple options and, if they selected "other", they were asked to specify.

Participants self-reported that, on average, they spent 47% of their time engaging in virtual social interactions during the iso media session (std = 28%; range = 5-100%). Most participants (35 out of 40) engaged with virtual social interactions more than 20% of iso media session and 18 out of 40 participants (45%) spent more than 50% of the session engaging in virtual social interactions. We asked participants what they used for connecting with others (multiple selections possible) and the majority of

participants reported using instant messaging (37 out of 40). Some participants reported posting (9) and a small number reported engaging in voice calls (3), video calls (2) or gaming (3). Participants mainly used Snapchat (28), Instagram (27) and WhatsApp (23) to connect with others, followed by TikTok (10), Twitter (3), Discord (3) and Facebook/Facebook Messenger (2). The majority of participants connected virtually with friends (38), followed by family (19), romantic partners (13) and acquaintances (4).

Behavioural: Reward tasks

We used two computerized behavioural tasks to measure reward responsiveness in two domains (i.e., reward seeking and reward learning) at baseline and at the end of the two isolation sessions (iso total and iso media).

Reward seeking. We used an effort-based decision making (EBDM) task, which was an adapted version of the Effort Expenditure for Rewards Task (EEfRT[65]). In our adapted version, we manipulated and orthogonalized reward and effort. Participants were presented with a series of trials with different combinations of high rewards and low rewards for successfully completing “hard tasks” and “easy tasks”. First, participants underwent two calibration trials in which the task measured their maximum effort (the maximum number of button presses performed in 10 s using their index finger). Based on this, “hard task” trials required participants to make 80% of the maximal button presses they managed in 10 s to succeed. “Easy task” trials required participants to make 40% of the maximal button presses they managed in 10 s to succeed. Rewards earned could be either 1 point (low reward) or 4 points (high reward) and participants were presented with combinations of required effort (hard/easy) and potential rewards (high/low) on each trial before deciding whether they wanted to perform that trial’s task. Points were converted into monetary value at the end of each session.

In addition, we added the element of context (nature vs social). After successfully performing a trial, participants saw a message indicating how much they won (1 point or 4 points depending on the reward value of that trial) and pictures of landscapes (nature context; e.g., pleasant pictures of mountains, beaches, rivers, etc.) or social pictures (social context; e.g., pleasant pictures of people hugging, laughing together, etc.). Participants either saw 1 picture or 4 pictures, corresponding to the number of points they earned in that trial (i.e., 1 picture for 1 point and 4 pictures for 4 points). This allowed us to measure the extent to which effort-based decision-making was modulated by reward and whether isolation selectively enhanced reward seeking in social vs non-social contexts. In total, there were 8 conditions: high effort - high reward; low effort – high reward; high effort – low reward; low effort - low reward, each in either a nature or social context, with 6 trials per condition (48 trials in total), which were presented in random order.

Participants had to indicate whether or not they wanted to play each trial. There was no time restriction for participants’ decisions. If participants completed the trial successfully (i.e., pumped up the bar within the time limit), they saw the social or nature pictures described above. Pictures were presented for 8 s

together with the number of points participants won. After each trial, a fixation cross was presented for 0.5 s.

Reward Learning. The ability to learn stimulus–reinforcement associations and to reverse them based on probabilistic feedback was measured using a probabilistic reinforcement and reversal learning task[66,67]. In the present task, participants were shown two slot machines and asked to choose between them to obtain a reward. One slot machine was rewarded 80% of the time; the other was rewarded 20% of the time. Participants needed to learn from feedback through trial and error which slot machine was rewarded more often. After 7 trials, the reward contingencies switched and participants needed to learn the new reward contingencies. Feedback was given via symbols (non-social feedback) and facial expressions (social feedback) in two counterbalanced blocks (28 trials per block, in total 56 trials). Rewards (positive feedback) were represented by either a plus symbol (+; non-social) or a smiling face (social), while the absence of a reward (negative feedback) was represented by a zero symbol (0; non-social) or a neutral face (social).

The two facial pictures (smiling and neutral) were generated by averaging facial pictures of Caucasian young adults. To do this, we used happy and neutral faces from the Averaged Karolinska Directed Emotional Faces set[68]. We averaged the female and male faces for each emotion (happy and neutral) rendering them ambiguous as to gender. Photoshop was used to create the averages. We cropped the images to remove the background/hair and display just the face. Facial pictures were displayed in black and white to match the non-social feedback. Participants were given 1 s to respond on each trial and then received feedback for 0.5 s. A fixation cross was presented for 0.5 s between each trial.

Participants also underwent three other tasks during the experiment (a go-no go task, a peer influence task and a threat learning task and (see full description of experimental procedures here: https://osf.io/kbgsv*).* The order of the tasks was counterbalanced between participants with the exception that the threat learning task was always presented after the reward tasks reported here, so that there would be no effect of threat exposure on reward processing. Results from the other behavioural tasks will be reported elsewhere.

Data analysis

Reward Seeking. We calculated the sum of number of played trials in the EBDM task and mean response times across all trials for the following conditions (nature or social): high effort - high reward; low effort – high reward; high effort – low reward; low effort - low reward (8 conditions in total). Response times (RTs) to choose whether to undergo an effortful task have been shown to be indicative of strength of preference[27] and here were used as the main measure of reward seeking due to the low variance in participants' choice data (most participants chose to play every trial in the task). We assessed whether the number of played trials and RTs when deciding to play differed between sessions (baseline, iso total, and iso media). We used mixed effects models to investigate fixed effects of session, context, reward and

effort on choices and RTs (for trials in which participants decided to do the effortful task) using separate models for choices and RTs. Subject was included as a random effect (allowing intercepts and slopes to vary between participants[69]). Calibration (i.e., the number of button presses participants managed at the beginning of each session from which hard and easy tasks were calculated) was added as a control variable in the model.

The command for both models (i.e., choice and RT data) was: *fitlme(Data,'Response~ (session*effort*reward*context)+(Calibration)+(session*effort*reward*context|subjectID)').*

Reward Learning. Participants' choices from the RL task were analysed using a computational reinforcement learning and decision-making model for probabilistic reversal learning tasks[66].

To assess learning during the task, we first extracted the trial-by-trial responses for each participant, then employed a computational reinforcement learning model for probabilistic reversal learning using a hierarchical Bayesian modelling approach[66] to estimate the learning rates for each participant.

Here, individual parameters are drawn from group-level normal distributions. We used standard normal priors for the group-level means[e.g., 70–72]. For the group-level standard deviations, we used half-Cauchy prior distributions, which tend to give sharper and more reasonable estimates than uniform or inverse-Gaussian prior distributions (see reference for details [73]).

To select a model that best captured the behaviour of participants, we first compared model fits between three different reversal learning models on the data from each session. Each model employs the Rescorla-Wagner value update rule[30] but differs in terms of how information is integrated. The Rescorla-Wagner update rule assumes that individuals assign and update internal stimulus value signals based on the prediction error, i.e., the mismatch between outcome (received reward/punishment following choice of this stimulus) and prediction (expected value of choosing this stimulus). The following models were compared:

(i) A reward-punishment model[29], which expands the classic Rescorla-Wagner model of conditioning[30] with separate learning rates for reward and punishment trials, here treating non-wins as punishments:

$$V_{c,t} = \begin{cases} V_{c,t-1} + \eta^{rew}(O_{t-1} - V_{c,t-1}), if O_{t-1} > 0 \\ V_{c,t-1} + \eta^{nrew}(O_{t-1} - V_{c,t-1}), if O_{t-1} < 0 \end{cases}$$

1

Where, ηrew is the learning rate for rewards and ηnrew is the learning rate for non-rewards; *O* is the outcome received. In this model, only the chosen stimulus value is updated. *Vc,t* is the value of choice *c* on trial *t. O > 0* indicates a win and *O < 0* indicates no win.

(ii) An experience-weighted attraction model[31], which includes an "experience" weight parameter that decouples acquisition and reversal by allowing the balance between past experience and new

information to increasingly tip in favour of past experience:

$$n_{c,t} = n_{c,t-1} \times \rho + 1$$

2

$$V_{c,t} = (V_{c,t-1} \times \phi \times n_{c,t-1} + O_{t-1})/n_{c,t}$$

3

Where, *nc,t* is the "experience weight" of the chosen stimulus on trial *t*, which is updated on every trial using the experience decay factor *ρ. Vc,t* is the value of choice *c* on trial *t* for outcome *O* received in response to that choice, and *φ* is the decay factor for the previous payoffs. In this model, *φ* is equivalent to the inverse of the learning rate in Rescorla-Wagner models.

(iii) A fictitious update model[33], which includes an update rule for the unchosen option considering the knowledge individuals gain about the unchosen option, here with separate learning rates for positive and negative prediction errors.

$$V_{c,t} = \begin{cases} V_{c,t-1} + \eta^{rew} \left(O_{t-1} - V_{c,t-1}\right), if O_{t-1} > 0 \\ V_{c,t-1} + \eta^{nrew} \left(O_{t-1} - V_{c,t-1}\right), if O_{t-1} < 0 \end{cases}$$

4

$$V_{nc,t} = \begin{cases} V_{nc,t-1} + \eta^{nrew} \left(-O_{t-1} - V_{nc,t-1}\right), if O_{t-1} > 0 \\ V_{nc,t-1} + \eta^{rew} \left(-O_{t-1} - V_{nc,t-1}\right), if O_{t-1} < 0 \end{cases}$$

5

Where the value *V* of both the chosen *c* and unchosen *nc* stimulus are updated with the actual prediction error and the counterfactual prediction error per trial *t*, respectively. *O* is the outcome received. The learning rate η, which is divided into *η*rew (learning rate for rewards; learning_pos) and *η*nrew (learning rate for non-rewards; learning_neg) evidences the magnitude of the value update affected by both positive and negative prediction errors.

Model fitting. For all models, a softmax choice function was used to compute the action probability given the action values. On each trial *t*, the action probability of choosing option A (over B) was defined as follows:

$$p\left(A\right) = \frac{1}{1 + e^{-\beta(V_A - V_B)}}$$

6

Where β is the inverse temperature parameter that governs the stochasticity of the choice, computed using inverse logit transfer. Higher β values denote decisions driven by relative value whereas lower β values denote more choice stochasticity.

Parameter estimation was performed with hierarchical Bayesian analysis using Stan language in R via the hBayesDM package[66]. Posterior inference was performed using Markov Chain Monte Carlo (MCMC) sampling using the mode of the posterior distribution as the summary parameter for individual participants. Four MCMC chains were used, with 1000 post-warmup iterations per chain, resulting in 4000 valid MCMC samples. Model convergence was assessed by examining R-hat values, an index of the convergence of the chains[74]. R-hat values of all models were lower than 1.1 suggesting MCMC samples were well mixed and converged to stationary distributions.

The models were fitted separately to the data from each session (baseline, iso total and iso media) and condition (social, non-social) to assess whether model fit would be affected by session or condition.

Model comparison. Comparison of model fit was assessed using Bayesian bootstrap and model averaging, whereby log-likelihoods for each model were evaluated at the posterior simulations and a weight obtained for each model. We compared Bayesian model weights between models (higher weights denoting better model fit). Model parameters were then extracted for each model that provided the best estimate within a data set (here defined as > 0.50 model weight). We found that model fit was best for the fictitious update model for each data set (Table S5).

Following model selection, we extracted individual participants' learning rates and inverse temperature parameters from the best performing model (i.e., the fictitious update model) for each session (baseline, iso total and iso media) and condition (social, non-social).

To compare differences in model parameters, we used mixed effects models to test for differences between sessions to estimate the fixed effects of feedback condition (social vs non-social) and session (baseline, iso total, and iso media) on three parameters of interest: learning_pos (learning rate from positive prediction errors (PEs), learning_neg (learning rate from negative PEs) and beta (inverse temperature). Separate models were calculated for each parameter. Subject was included as a random effect (allowing intercepts and slopes to vary between participants[69]).

The command for each model was:

*fitlme(Data,'Parameter ~ (session * condition) + (session * condition|subjectID)').*

Simulations. Whether a certain difference in parameters indicates "better" or "worse" behaviour can be heavily dependent on the task design and the values of the other relevant model parameters[37]. We therefore used data simulations to identify the optimal learning parameters for the present task environment to help with interpretation of the results. The parameter combinations were taken from a grid spanned by the learning rate (η; minimum value/maximum value/steps = 0/1/60) and "inverse temperature β" (0/20/60). Each of these virtual participants then completed 28 trials corresponding to the actual number of trials in our task for each condition (social, non-social) in each session (baseline, iso total, iso media) with a reward schedule of 80:20. For each virtual participant, we then calculated the percentage of correct choices as the percentage of choices for the option that was associated with a

higher probability for a reward. To reduce random noise due to the finite number of samples, we smoothed the resulting images with a gaussian filter (std = 2). We identified the optimal learning rate and inverse temperature as the values that resulted in the highest choice accuracy (average of 10 highest accuracy scores in the simulations; Fig. S2).

Data analysis and visualization for the reward tasks was implemented in Python (version: 3.7) using Jupyter notebooks (using packages Pandas, NumPy and Seaborn), RStudio and Matlab2020a.

MRI Data Analysis

Preprocessing. We used open source preprocessing pipelines for fMRI data, developed through the nipy and nipype initiatives[75]. We used the heudiconv python application which uses dcm2niix to convert raw scanner data into the NIFTI image file format, then organize that data into a BIDS-formatted directory structure. The fMRIprep application[76] was used to minimally preprocess the anatomical and functional data (using default settings but including susceptibility distortion correction using fieldmaps (see below).

Because we collected fMRI data on 7T and used an MP2RAGE sequence, which is not a format that is currently supported by fMRIprep, we first processed the MP2RAGE data so it could be used by fMRIprep. To this end, we applied an offline reconstruction method using in-house code that reconstructs a phase-sensitive inversion recovery (PSIR) image from MP2RAGE scans (nifti files). The PSIR file was then added as a T1 file into fMRIprep to be processed and was treated as a T1 anatomical scan in fMRIprep.

Using fMRIprep, we then skull-stripped anatomical images, first roughly using the atlas-based ANTS program[77], and refined this using information from FreeSurfer surfaces after reconstruction was completed[78]. Brain tissue segmentation was performed with the FMRIB Software Library (FSL) FAST program[79]. Images were spatially normalized to MNI-space using the multiscale, mutual-information based, nonlinear registration scheme implemented in ANTS. We visually inspected brain masks, tissue segmentation and FreeSurfer surfaces. Susceptibility distortion correction was performed using phase-difference B0 estimation[80].

A reference image for each run was generated from the input BOLD timeseries. A functional brain mask was created using a combination of FSL, ANTS, AFNI and nilearn tools[81]. Using FSL's MCFLIRT program[82] we estimated and corrected for head motion, resulting in a coregistered BOLD series as well as motion-based confound regressors. Any run containing a framewise displacement greater than 0.4 mm on more than 25% of the total frames was excluded from additional analyses. Additional confound regressors were generated, including other measures of motion (framewise displacement and DVARS and anatomical CompCor[83] timeseries derived from cerebrospinal fluid (CSF) and white matter tissue segments. The reference image of each run was aligned with the anatomical image using FreeSurfer's program "bbregister"[84]. The timepoint-to-functional reference transformation, the functional reference to

anatomical transformation, and the anatomical-to-MNI transformation were concatenated into a single transformation and used to transform each functional timeseries into MNI template space.

Spatial smoothing was performed on the fMRIprep outputs with a 6 mm smoothing kernel using FSL's SUSAN tool[85], which uses segmentation boundaries to avoid smoothing across tissue types. MRIQC, an open-source quality assurance software tool[86], was used to generate additional reports which display Image Quality Metrics (IQMs).

Functional MRI Data Modelling. Analyses were conducted using the nipype framework[75]. For run-level analyses, the preprocessed timeseries were assessed with algorithms from the Artifact Removal Toolbox (ART) to identify frames within the run that have an abnormal amount of motion (0.4 mm of total displacement, or an intensity spike greater than 3 standard deviations from mean). The design matrix included boxcars for the experimental conditions convolved with a double-gamma hemodynamic response function (HRF), and nuisance regressors representing frame-wise motion, the anatomical CompCor regressors derived from white matter and CSF, as well as impulse regressors for volumes identified by ART. A high-pass filter was applied to the design matrix and the smoothed data. The model was evaluated using FSL's FILM[87]. Subject-level contrast maps were generated using FSL's FLAME[87] in mixed-effects mode.

ROIs in MID task. To analyse the functional data from the MID task, we applied whole brain analyses and region of interest (ROI) analyses. Whole brain data was used to generate group-level contrasts for the high reward vs no reward conditions to assess whether the resulting group-level brain activation was in line with expected activation for this contrast[38]. For the ROI analyses, we included regions that have been consistently activated during reward anticipation in an Activation Likelihood Estimation meta-analysis of 50 fMRI studies using the MID task (Oldham 2018[38]). These regions comprised: the ventral striatum (VS), midbrain (MB), amygdala (AMY), anterior insula (AI), occipital cortex (OC), thalamus (TH) and supplementary motor area (SMA). ROIs were created using the Harvard–Oxford cortical and subcortical probabilistic anatomical atlases; included in FSL. We were not able to find an anatomical mask for the anterior insula and therefore created ROIs based on peak activation locations reported in Oldham 2018 (left AI: -40 14 − 8; right AI: 34 24 − 2). We drew 5 mm spheres around these peak coordinates and created a binary mask combining left and right AI.

Quantifying neural reward sensitivity.

Univariate. We extracted mean activation from the contrast high reward > no reward during the anticipation phase of the MID task as a measure of univariate neural reward sensitivity from each ROI (7 ROIs in total).

Multivariate. Multivariate analyses can detect differences between conditions with higher sensitivity than conventional univariate analyses[88]. We therefore also employed a multivariate measure of neural reward sensitivity by quantifying the similarity of the multivariate spatial patterns of activity between high

reward versus no reward during the anticipation phase of the MID task from each ROI. Multivariate classification accuracy here can be seen as an indication of the amount of information about a variable of interest available in the BOLD signal[89,90], such that individual differences in classification accuracy indicate neural reward sensitivity in the multivariate domain. However, differences between multivariate and univariate tests do not afford conclusions about the nature or dimensionality of the neural signal (as both can also stem from the same source[88]). Thus, here, these measures were used to assess the same concept (i.e., neural reward sensitivity) and we corrected for multiple comparisons across the different measures. For each ROI, we extracted the β values from the generalized linear model (that is, the amplitude of the fit HRF) of the response to the two conditions of interest (high reward, no reward) for each trial in each run (28 trials per condition per run × 4 runs) resulting in 112 β values for each voxel. Responses were extracted from all voxels in the anatomically identified ROIs in each participant (see *ROIs in MID task* for ROI definition). No additional feature selection was applied. All multivariate analyses were conducted with the PyMVPA[91] toolbox in Python (http://www.pymvpa.org) and Matlab 2020a. We first trained a machine learning algorithm to decode neural patterns of high reward vs no reward for each subject within each ROI. We used a 4-fold linear support vector machine (SVM; using linear kernel) classification for which we trained the classifier on 3 of the runs from the MID task while one run was left out for testing. We did this 4 times (4-folds) so that each run was a testing data set on one of the folds. We then averaged classification accuracies from each fold to create a measure of classification accuracy for each ROI for each participant. To obtain confidence intervals of the mean classification accuracy for each ROI, we used bootstrapping. We generated 1,000 datasets randomly by sampling with replacement from the classification accuracies for each ROI across participants using Matlab's bootci function.

First, we tested whether these classification accuracies for each ROI across participants were significantly above chance, indicating the ability to decode high reward vs no reward representations in that ROI. While the selected ROIs were chosen based on highly consistent univariate responsiveness to the MID task across different studies, they might not necessarily also represent reward anticipation in the multivariate domain. For hypothesis testing at the group level, we used a permutation analysis following the methods in Stelzer et al.[92] (also described in Tomova et al.[24]). This nonparametric approach does not depend on assumptions about the distribution of classification accuracies[92,93]. To generate a null distribution from the data, we followed the steps described in Stelzer et al., which we summarize below. We shuffled the condition labels randomly during training within each run, and then used the same 4-fold cross-validation approach as in the original analysis to obtain prediction accuracies for each fold, which we then averaged to create a measure of classification accuracy for each ROI for each participant. We performed this permutation analysis 100 times per participant (thus creating 100 random permutations), resulting in 100 accuracy values per participant. To create a null distribution at the group level, we then randomly drew one of the 100 accuracy values for each participant and calculated a mean across participants. This procedure was repeated 10^5 times for each ROI, creating the null distributions for each ROI. We calculated the probability P of obtaining a mean accuracy value in the null distributions that is equal to or higher than the true mean from the analyses. Following Stelzer et al., we rejected the null hypothesis of no group-level decoding if $P < 0.001$.

We were able to decode high reward vs no reward representations in all of our 7 selected ROIs: VS (mean accuracy = 0.568, bootstrapped CI = 0.540,0.594, P < 0.001); MB (mean accuracy = 0.551, bootstrapped CI = 0.525,0.582, P < 0.001); OC (mean accuracy = 0.579, bootstrapped CI = 0.552,0.606, P < 0.001); AI (mean accuracy = 0.545, bootstrapped CI = 0.521,0.573, P < 0.001); TH (mean accuracy = 0.560, bootstrapped CI = 0.534,0.592, P < 0.001); SMA (mean accuracy = 0.568, bootstrapped CI = 0.536,0.601, P < 0.001); AMY (mean accuracy = 0.506, bootstrapped CI = 0.482, 0.534, P < 0.001).

We used the classification accuracy obtained from each participant from the SVM classification as a measure of multivariate neural reward sensitivity for subsequent analyses.

For both measures of neural reward sensitivity (univariate and multivariate), we used Pearson correlations to test for an association between neural reward sensitivity and measures of reward seeking (RTs from the high reward / high effort conditions (mean across nature and social contexts) in the EBDM task) and reward learning (learning_pos rates in the RL task) in the isolation session. We calculated separate correlations for each ROI for each measure of neural reward sensitivity (univariate and multivariate; 14 tests in total) and report results as significant at p < 0.003 (0.05/14)). ROI data analysis and visualization was implemented in Python (version: 3.7) using Jupyter notebooks (using packages PyMVPA, SciPy, Pandas, NumPy and Seaborn) and Matlab2020a.

Declarations

Acknowledgments

This research was carried out at the Department of Psychology and the Wolfson Brain Imaging Centre at the University of Cambridge. The authors thank Michele Ferraro, Vicky Lupson and Tracy Horn for their support with data collection; Lei Zhang, Stefano Palminteri and Giacomo Bignardi for discussions on methods; and Olivier Mougin, Catarina Rua and the UK7T Network (*http://www.uk7t.org*) for developing the original PSIR conversion script.

Funding: We gratefully acknowledge support of this project by a Henslow Research Fellowship from the Cambridge Philosophical Society (to L.T.). SJB is funded by Wellcome (grant number WT107496/Z/15/Z), the MRC, the Jacobs Foundation, the Wellspring Foundation and the University of Cambridge. E.T. is funded by a Gates Cambridge Scholarship.

Author contributions: L.T. and S-J.B. designed the study with input from E.T. and K.T.; L.T., E.T. and K.T. collected the data; L.T. analyzed the data with support from S-J.B. and E.T.; L.T. wrote the manuscript and all authors provided feedback on the final version.

Competing interests: The authors declare no competing interests.

Data and materials availability: All de-identified neuroimaging data will be publicly available on

https://openneuro.org. Behavioural data will be made publicly available on the open science framework and analysis scripts and code used to generate the figures will be submitted to a publicly available Github repository. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) license to any Author Accepted Manuscript version arising from this submission.

References

1. Cacioppo, J. T., Cacioppo, S. & Boomsma, D. I. Evolutionary mechanisms for loneliness. Cogn. Emot. 28, 3–21 (2014).
2. Baumeister, R. F. & Leary, M. R. The need to belong: desire for interpersonal attachments as a fundamental human motivation. Psychol. Bull. 117, 497–529 (1995).
3. Twenge, J. M., Spitzberg, B. H. & Campbell, W. K. Less in-person social interaction with peers among US adolescents in the 21st century and links to loneliness. J. Soc. Pers. Relat. 36, 1892–1913 (2019).
4. Twenge, J. M. *et al.* Worldwide increases in adolescent loneliness. J. Adolesc. 93, 257–269 (2021).
5. Hammond, C. Who feels lonely? The results of the world's largest loneliness study. BBC https://www.bbc.co.uk/programmes/articles/2yzhfv4DvqVp5nZyxBD8G23/who-feels-lonely-the-results-of-the-world-s-largest-loneliness-study (2019).
6. Ladd, G. W. & Ettekal, I. Peer-related loneliness across early to late adolescence: normative trends, intra-individual trajectories, and links with depressive symptoms. J. Adolesc. 36, 1269–1282 (2013).
7. Qualter, P. *et al.* Trajectories of loneliness during childhood and adolescence: Predictors and health outcomes. J. Adolesc. 36, 1283–1293 (2013).
8. Prinstein, M. J. & La Greca, A. M. Peer crowd affiliation and internalizing distress in childhood and adolescence: a longitudinal follow-back study. J. Res. Adolesc. 12, 325–351 (2002).
9. Goosby, B. J., Bellatorre, A., Walsemann, K. M. & Cheadle, J. E. Adolescent Loneliness and Health in Early Adulthood. Sociol. Inq. 83, 10.1111/soin.12018 (2013).
10. Hawkley, L. C. & Capitanio, J. P. Perceived social isolation, evolutionary fitness and health outcomes: a lifespan approach. Philos. Trans. Biol. Sci. 370, 1–12 (2015).
11. Essau, C. A., de la Torre-Luque, A., Lewinsohn, P. M. & Rohde, P. Patterns, predictors, and outcome of the trajectories of depressive symptoms from adolescence to adulthood. Depress. Anxiety (2020).
12. Kayaoğlu, K., Okanlı, A., Budak, F. K. & Aslanoğlu, E. The correlation between loneliness and substance use proclivity in child and adolescent substance users. J. Subst. Use 27, 70–73 (2022).
13. Dogan-Sander, E., Kohls, E., Baldofski, S. & Rummel-Kluge, C. More Depressive Symptoms, Alcohol and Drug Consumption: Increase in Mental Health Symptoms Among University Students After One Year of the COVID-19 Pandemic. Front. Psychiatry 12, (2021).
14. Bonar, E. E. *et al.* Longitudinal within- and between-person associations of substance use, social influences, and loneliness among adolescents and emerging adults who use drugs. Addict. Res. Theory 30, 262–267 (2022).

15. Ozeylem, F., de la Torre-Luque, A. & Essau, C. A. Factors related to substance use among adolescents from six low-and middle-income countries. Addict. Behav. Reports 14, (2021).
16. Hall, F. S. Social deprivation of neonatal, adolescent, and adult rats has distinct neurochemical and behavioral consequences. Crit. Rev. Neurobiol. 12, 129–162 (1998).
17. Novick, A. M. *et al.* The effects of early life stress on reward processing. J. Psychiatr. Res. 101, 80–103 (2018).
18. Bayindir, N. & Paisley, E. W. Digital vs Traditional Media Consumption. globalwebindex vol. Trend Repo (2019).
19. Nowland, R., Necka, E. A. & Cacioppo, J. T. Loneliness and Social Internet Use: Pathways to Reconnection in a Digital World? Perspect Psychol Sci 13, 70–87 (2018).
20. Kraut, R. *et al.* Internet paradox. A social technology that reduces social involvement and psychological well-being? Am Psychol 53, 1017–1031 (1998).
21. Nie, N. H. Sociability, Interpersonal Relations, and the Internet:Reconciling Conflicting Findings. Am. Behav. Sci. 45, 420–435 (2001).
22. Alison Bryant, J., Sanders-Jackson, A. & Smallwood, A. IMing, Text Messaging, and Adolescent Social Networks. J. Comput. Commun. 11, 577–592 (2006).
23. Valkenburg, P. M. & Peter, J. Preadolescents' and adolescents' online communication and their closeness to friends. Dev Psychol 43, 267–277 (2007).
24. Tomova Wang, K., Thompson, T., Matthews, G., Takahashi, A., Tye, K., Saxe, R., L. Acute social isolation causes midbrain craving responses similar to hunger. Nat. Neurosci. 23, 1597–1605 (2020).
25. Lasgaard, M. Reliability and validity of the Danish version of the UCLA Loneliness Scale. Pers. Individ. Dif. 42, 1359–1366 (2007).
26. Russell, D. W. UCLA Loneliness Scale (Version 3): reliability, validity, and factor structure. J Pers Assess 66, 20–40 (1996).
27. Konovalov, A. & Krajbich, I. Revealed strength of preference: Inference from response times. Judgm. Decis. Mak. 14, 381–394 (2019).
28. Palminteri, S., Lefebvre, G., Kilford, E. J. & Blakemore, S.-J. Confirmation bias in human reinforcement learning: Evidence from counterfactual feedback processing. PLoS Comput. Biol. 13, e1005684 (2017).
29. den Ouden, H. E. M. *et al.* Dissociable effects of dopamine and serotonin on reversal learning. Neuron 80, 1090–1100 (2013).
30. Rescorla, R. A. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. Curr. Res. theory 64–99 (1972).
31. Camerer, C. & Hua Ho, T. Experience-weighted Attraction Learning in Normal Form Games. Econometrica 67, 827–874 (1999).
32. Katahira, K. The statistical structures of reinforcement learning with asymmetric value updates. J. Math. Psychol. 87, 31–45 (2018).

33. Gläscher, J., Hampton, A. N. & O'Doherty, J. P. Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. Cereb. Cortex 19, 483–495 (2009).
34. Hampton, A. N., Bossaerts, P. & O'doherty, J. P. The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. J. Neurosci. 26, 8360–8367 (2006).
35. Hampton, A. N., Adolphs, R., Tyszka, J. M. & O'Doherty, J. P. Contributions of the amygdala to reward expectancy and choice signals in human prefrontal cortex. Neuron 55, 545–555 (2007).
36. Crawley, D. *et al.* Modeling flexible behavior in childhood to adulthood shows age-dependent learning mechanisms and less optimal learning in autism in each age group. PLoS Biol. 18, e3000908 (2020).
37. Zhang, L., Lengersdorff, L., Mikus, N., Gläscher, J. & Lamm, C. Using reinforcement learning models in social neuroscience: frameworks, pitfalls and suggestions of best practices. Soc. Cogn. Affect. Neurosci. 15, 695–707 (2020).
38. Oldham, S. *et al.* The anticipation and outcome phases of reward and loss processing: A neuroimaging meta-analysis of the monetary incentive delay task. Hum. Brain Mapp. 39, 3398–3418 (2018).
39. Shadmehr, R., Huang, H. J. & Ahmed, A. A. A Representation of Effort in Decision-Making and Motor Control. Curr. Biol. 26, 1929–1934 (2016).
40. Matthews, G. A. *et al.* Dorsal Raphe Dopamine Neurons Represent the Experience of Social Isolation. Cell 164, 617–631 (2016).
41. McCool, B. A. & Chappell, A. M. Early social isolation in male Long-Evans rats alters both appetitive and consummatory behaviors expressed during operant ethanol self-administration. Alcohol. Clin. Exp. Res. 33, 273–282 (2009).
42. Amitai, N. *et al.* Isolation rearing effects on probabilistic learning and cognitive flexibility in rats. Cogn. Affect. Behav. Neurosci. 14, 388–406 (2014).
43. Schrijver, N. C. A., Pallier, P. N., Brown, V. J. & Würbel, H. Double dissociation of social and environmental stimulation on spatial learning and reversal learning in rats. Behav. Brain Res. 152, 307–314 (2004).
44. Lindström, B. *et al.* A computational reward learning account of social media engagement. Nat. Commun. 12, 1311 (2021).
45. Southwick, S. M., Vythilingam, M. & Charney, D. S. THE PSYCHOBIOLOGY OF DEPRESSION AND RESILIENCE TO STRESS: Implications for. Annu. Rev. Clin. Psychol 1, 255–291 (2005).
46. Vidal-Ribas, P. *et al.* Bidirectional Associations Between Stress and Reward Processing in Children and Adolescents: A Longitudinal Neuroimaging Study. Biol. psychiatry. Cogn. Neurosci. neuroimaging 4, 893–901 (2019).
47. Nikolova, Y. S., Bogdan, R., Brigidi, B. D. & Hariri, A. R. Ventral striatum reactivity to reward and recent life stress interact to predict positive affect. Biol. Psychiatry 72, 157–163 (2012).

48. Jauhar, S. *et al.* Brain activations associated with anticipation and delivery of monetary reward: A systematic review and meta-analysis of fMRI studies. PLoS One 16, e0255292 (2021).
49. Ress, D., Backus, B. T. & Heeger, D. J. Activity in primary visual cortex predicts performance in a visual detection task. Nat. Neurosci. 3, 940–945 (2000).
50. Chen, C. S., Knep, E., Han, A., Ebitz, R. B. & Grissom, N. M. Sex differences in learning from exploration. Elife 10, e69748 (2021).
51. Cacioppo, J. T. & Cacioppo, S. Chapter Three - Loneliness in the Modern Age: An Evolutionary Theory of Loneliness (ETL). in Advances in Experimental Social Psychology (ed. Olson, J. M.) vol. 58 127–197 (Academic Press, 2018).
52. Matthews, G. A. & Tye, K. M. Neural mechanisms of social homeostasis. Ann. N. Y. Acad. Sci. 1457, 5–25 (2019).
53. Steinberg, L. A Social Neuroscience Perspective on Adolescent Risk-Taking. Dev Rev 28, 78–106 (2008).
54. Berridge, K. C. & Robinson, T. E. Parsing reward. Trends Neurosci. 26, 507–513 (2003).
55. Burke, A. R., McCormick, C. M., Pellis, S. M. & Lukkes, J. L. Impact of adolescent social experiences on behavior and neural circuits implicated in mental illnesses. Neurosci. Biobehav. Rev. 76, 280–300 (2017).
56. Harris, P. A. *et al.* The REDCap consortium: Building an international community of software platform partners. J. Biomed. Inform. 95, 103208 (2019).
57. Shevlin, M., Murphy, S. & Murphy, J. The latent structure of loneliness: testing competing factor models of the UCLA loneliness scale in a large adolescent sample. Assessment 22, 208–215 (2015).
58. Von Der Heide, R., Vyas, G. & Olson, I. R. The social network-network: size is predicted by brain structure and function in the amygdala and paralimbic regions. Soc. Cogn. Affect. Neurosci. 9, 1962–1972 (2014).
59. Norbeck, J. S., Lindsey, A. M. & Carrieri, V. L. The development of an instrument to measure social support. Nurs Res 30, 264–269 (1981).
60. Chaudhari, A. J. *et al.* Spatial distortion correction and crystal identification for MRI-compatible position-sensitive avalanche photodiode-based PET scanners. IEEE Trans. Nucl. Sci. 56, 549–556 (2009).
61. Knutson, B., Westdorp, A., Kaiser, E. & Hommer, D. FMRI visualization of brain activity during a monetary incentive delay task. Neuroimage 12, 20–27 (2000).
62. Spielberger, C. D. State-trait anxiety inventory for adults. (1983).
63. Radloff, L. S. The CES-D scale: A self-report depression scale for research in the general population. Appl. Psychol. Meas. 1, 385–401 (1977).
64. Watson, D., Clark, L. A. & Tellegen, A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J. Pers. Soc. Psychol. 54, 1063 (1988).

65. Treadway, M. T., Buckholtz, J. W., Schwartzman, A. N., Lambert, W. E. & Zald, D. H. Worth the 'EEfRT'? The effort expenditure for rewards task as an objective measure of motivation and anhedonia. PLoS One 4, e6598 (2009).
66. Ahn, W.-Y., Haines, N. & Zhang, L. Revealing Neurocomputational Mechanisms of Reinforcement Learning and Decision-Making With the hBayesDM Package. Comput. Psychiatry 1, 24–57 (2017).
67. Metha, J. A. *et al.* Separating probability and reversal learning in a novel Probabilistic Reversal Learning task for mice. Front. Behav. Neurosci. 13, 270 (2020).
68. Lundqvist, D., & Litton, J. E. The Averaged Karolinska Directed Emotional Faces - AKDEF. CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet https://www.kdef.se/home/about akdef.html (1998).
69. Barr, D. J., Levy, R., Scheepers, C. & Tily, H. J. Random effects structure for confirmatory hypothesis testing: Keep it maximal. J Mem Lang 68, (2013).
70. Lee, M. D. How cognitive modeling can benefit from hierarchical Bayesian models. J. Math. Psychol. 55, 1–7 (2011).
71. Shiffrin, R. M., Lee, M. D., Kim, W. & Wagenmakers, E. A survey of model evaluation approaches with a tutorial on hierarchical Bayesian methods. Cogn. Sci. 32, 1248–1284 (2008).
72. Wetzels, R., Vandekerckhove, J., Tuerlinckx, F. & Wagenmakers, E.-J. Bayesian parameter estimation in the Expectancy Valence model of the Iowa gambling task. J. Math. Psychol. 54, 14–27 (2010).
73. Gelman, A. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). Bayesian Anal. 1, 515–534 (2006).
74. Gelman, A. & Rubin, D. B. Inference from Iterative Simulation Using Multiple Sequences. Stat. Sci. 7, 457–472 (1992).
75. Gorgolewski, K. et al. Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. Frontiers in Neuroinformatics vol. 5 (2011).
76. Esteban, O. *et al.* fMRIPrep: a robust preprocessing pipeline for functional MRI. Nat. Methods 16, 111–116 (2019).
77. Tustison, N. J. *et al.* Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements. Neuroimage 99, 166–179 (2014).
78. Dale, A. M., Fischl, B. & Sereno, M. I. Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage 9, 179–194 (1999).
79. Zhang, Y., Brady, M. & Smith, S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans. Med. Imaging 20, 45–57 (2001).
80. Jezzard, P. & Balaban, R. S. Correction for geometric distortion in echo planar images from B0 field variations. Magn. Reson. Med. 34, 65–73 (1995).
81. Abraham, A. et al. Machine learning for neuroimaging with scikit-learn. Frontiers in Neuroinformatics vol. 8 (2014).

82. Jenkinson, M., Bannister, P., Brady, M. & Smith, S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17, 825–841 (2002).
83. Behzadi, Y., Restom, K., Liau, J. & Liu, T. T. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. Neuroimage 37, 90–101 (2007).
84. Greve, D. N. & Fischl, B. Accurate and robust brain image alignment using boundary-based registration. Neuroimage 48, 63–72 (2009).
85. Smith, S. M. & Brady, J. M. SUSAN—A New Approach to Low Level Image Processing. Int. J. Comput. Vis. 23, 45–78 (1997).
86. Esteban, O. *et al.* MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. PLoS One 12, e0184661 (2017).
87. Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W. & Smith, S. M. FSL. Neuroimage 62, 782–790 (2012).
88. Davis, T. *et al.* What do differences between multi-voxel and univariate analysis mean? How subject-, voxel-, and trial-level variance impact fMRI analysis. Neuroimage 97, 271–283 (2014).
89. Kok, P., Jehee, J. F. M. & de Lange, F. P. Less Is More: Expectation Sharpens Representations in the Primary Visual Cortex. Neuron 75, 265–270 (2012).
90. Jehee, J. F. M., Brady, D. K. & Tong, F. Attention improves encoding of task-relevant features in the human visual cortex. J. Neurosci. 31, 8210–8219 (2011).
91. Hanke, M. *et al.* PyMVPA: a Python Toolbox for Multivariate Pattern Analysis of fMRI Data. Neuroinformatics 7, 37–53 (2009).
92. Stelzer, J., Chen, Y. & Turner, R. Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): random permutations and cluster size control. Neuroimage 65, 69–82 (2013).
93. Allefeld, C., Görgen, K. & Haynes, J.-D. Valid population inference for information-based imaging: From the second-level t-test to prevalence inference. Neuroimage 141, 378–392 (2016).

Figures

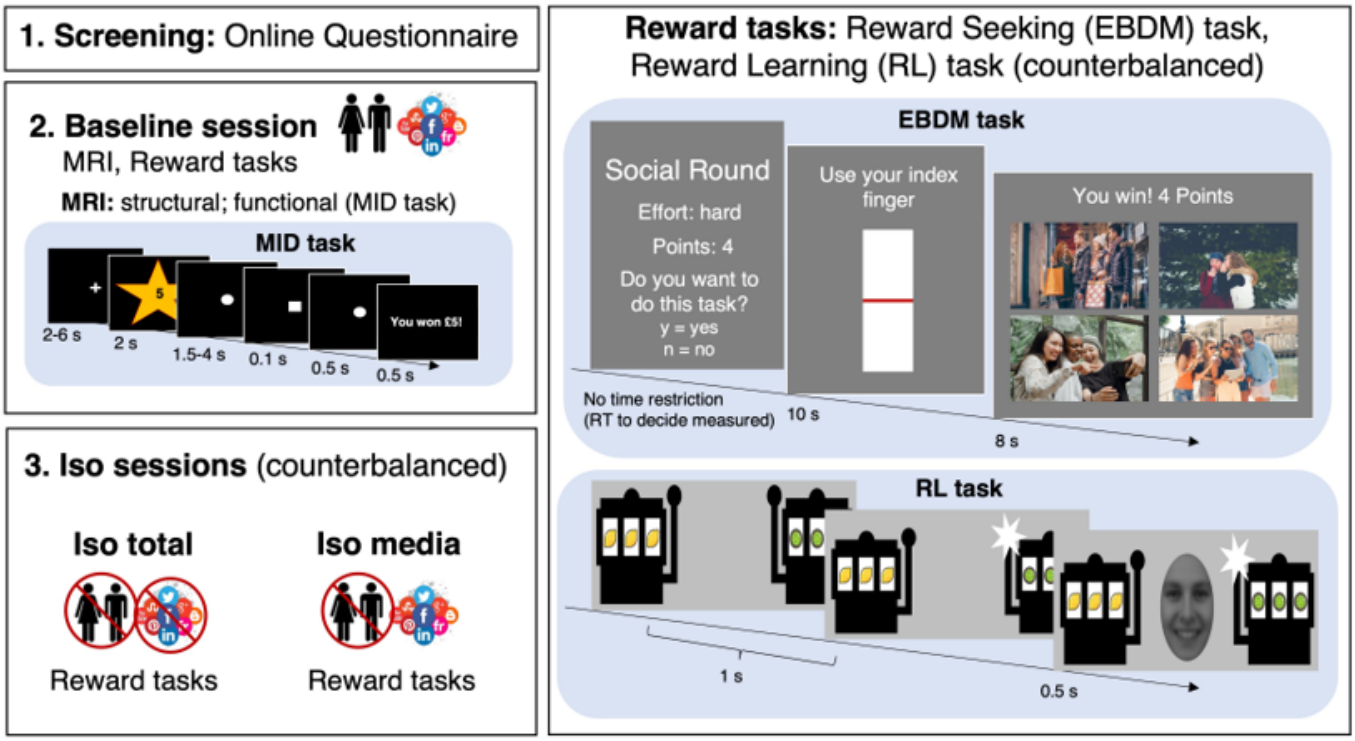

Figure 1

Overview of the experimental procedure

Left panel: Order of sessions. Individuals first completed an online screening questionnaire to assess eligibility and social connectedness. Eligible participants completed the baseline session, during which they underwent an MRI scan (functional MRI with a monetary incentive delay (MID) task). MID task: Participants were presented with a star indicating how much they could win on a given trial. Subsequently, they were asked to press a button as fast as possible after seeing a white square appear. Participants then saw a screen indicating how much (from 0 – £5) they had won on that trial. Following the scan, participants carried out the reward tasks (right panel). Participants were then invited to two isolation sessions (iso total, iso media; order counterbalanced). At the end of each of the three sessions, participants carried out two reward tasks. Right panel: Timelines of example trials for each reward task. Effort-based decision making (EBDM) task: On each trial, participants were first shown a screen that listed details for the current trial (level of effort (hard, easy), reward (high (4 points), low (1 point)) and context (social, nature)) and were asked whether they wanted to complete the trial. We measured the response time (RT) while participants decided if they wanted to complete the trial. If they chose yes, they saw a bar that they were required to pump up in 10 s (amount of effort required was calibrated to each participant's maximum effort levels assessed at the start of the task in each session) and, if they were successful, they saw a screen indicating how much they had won (1 or 4 points) and depicting images of social interactions or nature landscapes, depending on context (low reward = 1 image; high reward = 4 images). Reward Learning (RL) task: Participants were presented with two slot machines and asked to

choose one. After they chose, they received feedback as to whether they had won (indicated by a smiling face or a plus symbol) or did not win (indicated by a neutral face or a zero).

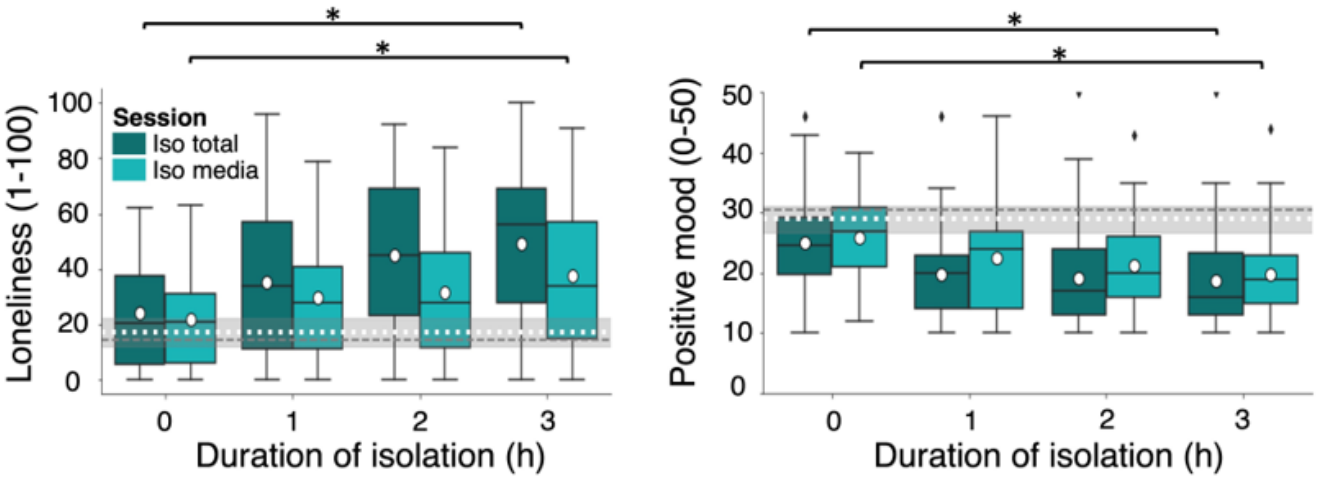

Figure 2

Effects of isolation on loneliness and mood.

Changes in self-reported loneliness (left) and positive mood (right) over time during isolation. Dark green = iso total session; turquoise = iso media session. The boxplots indicate the median (black centre line), mean (white circle), the interquartile range (IQR; box) and the 1.5 x IQR minima and maxima (whiskers). The black rhombi indicate outliers (values outside of 1.5 x IQR). The dashed lines across the plots indicate the mean (white) and median (grey) ratings during baseline and their 95% confidence interval (grey area). The asterisk indicates significance at $p < 0.001$.

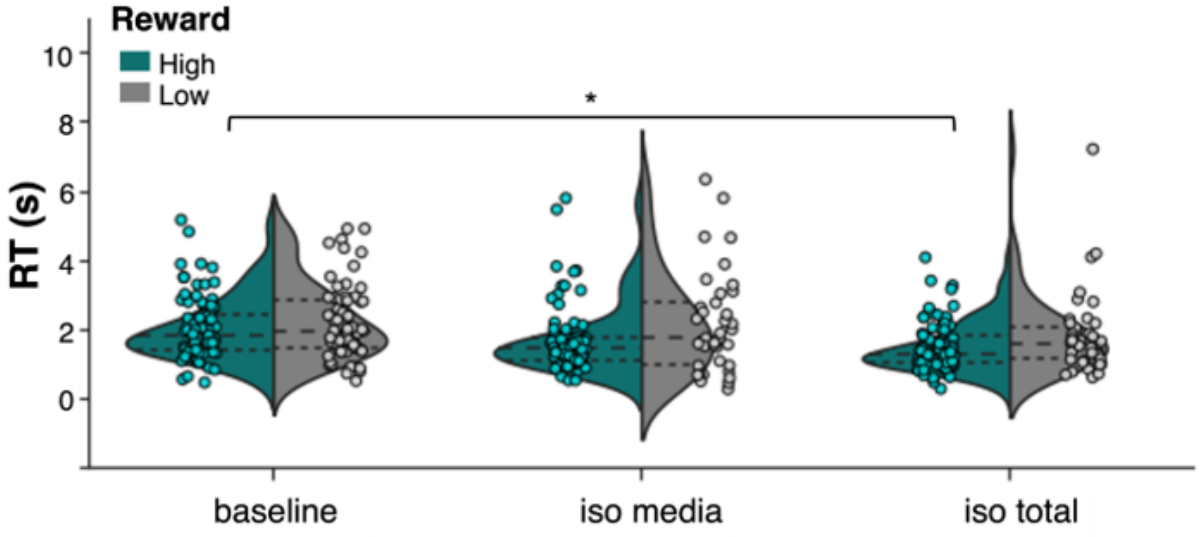

Figure 3

Effects of isolation on reward seeking.

Violin plots depict response times (RTs) during the reward seeking (EBDM) task for high effort trials in each session plotted separately for high rewards (green) and low rewards (grey) averaged across contexts (social, non-social). The violin plots illustrate the distribution of the data, the long-dashed lines indicate the median, the short-dashed lines the interquartile range with individual data points shown overlayed. The asterisk indicates significance at a Bonferroni-corrected p value < 0.0063.

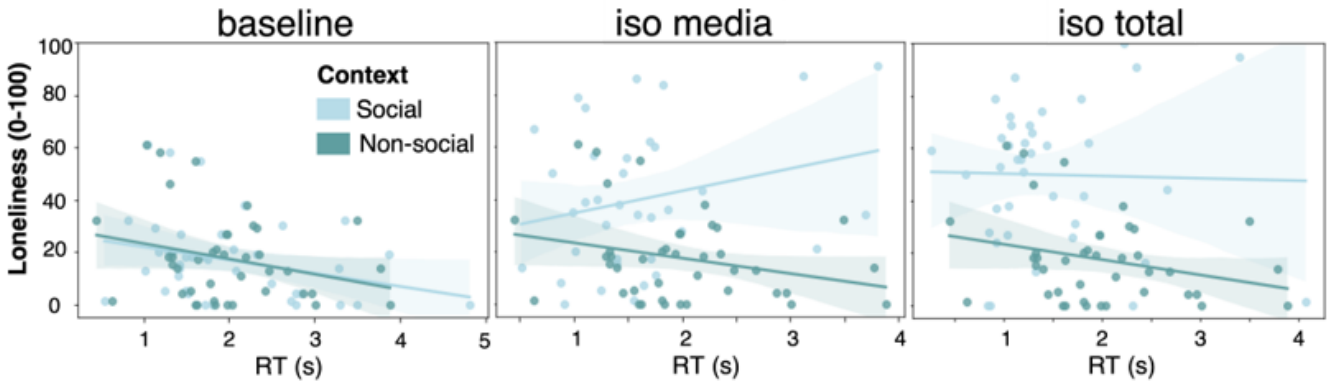

Figure 4

Correlations between self-reported loneliness after isolation and RTs in the EBDM task.

Scatter plots depict self-reported loneliness levels in each session (at the onset of the baseline session and after three hours of isolation in the two isolation sessions) and response times (RTs) during the EBDM task for high reward, high effort trials for the social and non-social context in each session (baseline, iso media, iso total). The scatter plots illustrate the main effect of loneliness on RTs in the EBDM task and the interaction between context and loneliness on RTs in the iso media session and show individual data points and regression lines indicating the linear fit for each context and the shaded areas depicts the 95% confidence interval. The plots show that, overall, higher loneliness in each session was associated with faster RTs in the task. However, in the iso media session in the social context this relationship was reversed and higher loneliness was associated with slower RTs in the task.

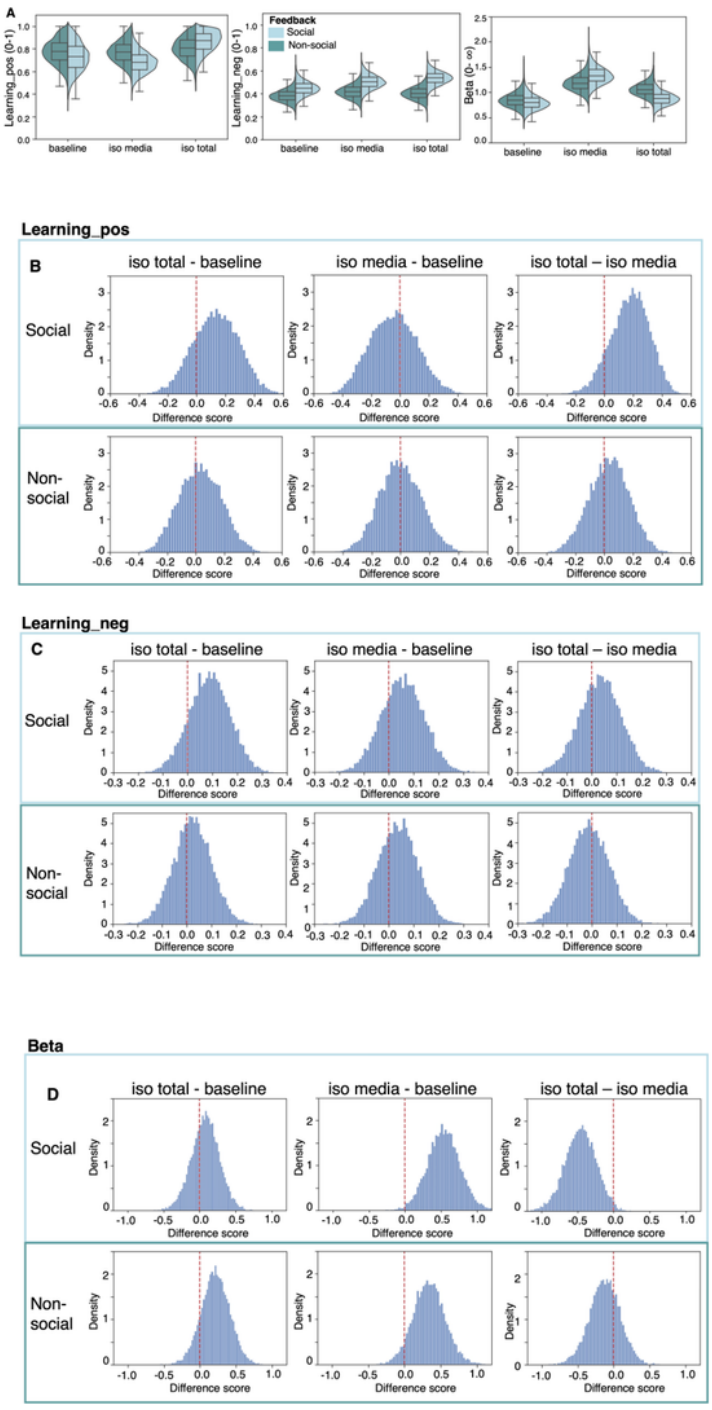

Figure 5

Fig. 6 Effects of isolation on reward learning.

A. Population-level mean posterior distributions of learning rates from positive prediction errors (learning_pos), learning rates from negative prediction errors (learning_neg) and inverse temperature (beta) for each session for each feedback type. The violin plots illustrate the posterior distributions for the

population-level mean effects of the different conditions. The posterior distributions consist of 12,000 samples combined from four Monte Carlo Markov chains. The boxplots illustrate the interquartile range (IQR; box) and the centre lines signify medians. The whiskers extend from the hinge to at most 1.5 times the interquartile range. B-D. Difference in population-level mean posterior distributions for each pair of sessions (iso total – baseline, iso media – baseline and iso total – iso media) and feedback type (social, non-social) for learning_pos rates (B), learning_neg rates (C) and beta (D). The histograms depict the density of difference scores; the red dashed line indicates zero (no difference).

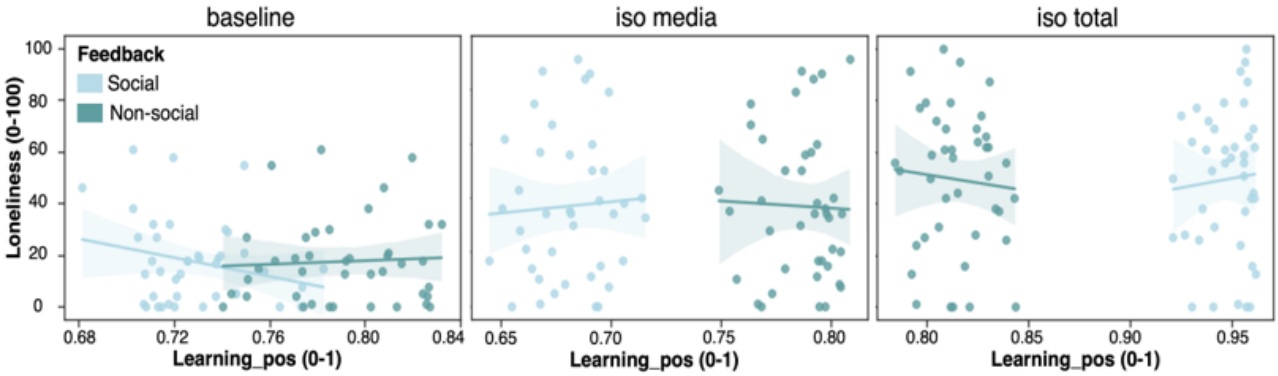

Figure 6

Fig. 7 Correlations between self-reported loneliness and learning_pos rates in the RL task.

Scatter plots depict self-reported loneliness levels in each session (at the onset of the baseline session and after three hours of isolation in the two isolation sessions) and learning_pos rates during the RL task for social (light green) and non-social (dark green) feedback for the baseline (left), iso media (middle) and iso total session (right). The scatter plots illustrate the interaction between loneliness, session and condition on learning_pos rates in the RL task. The scatter plots show the individual data points, the regression lines indicate the linear fit for each feedback type and the shaded areas depicts the 95% confidence interval. The plots show that in both isolation sessions (iso media and iso total) higher loneliness is associated with *higher* learning_pos rates from social feedback in the task but with *lower* learning_pos rates from non-social feedback. This relationship is reversed at baseline.

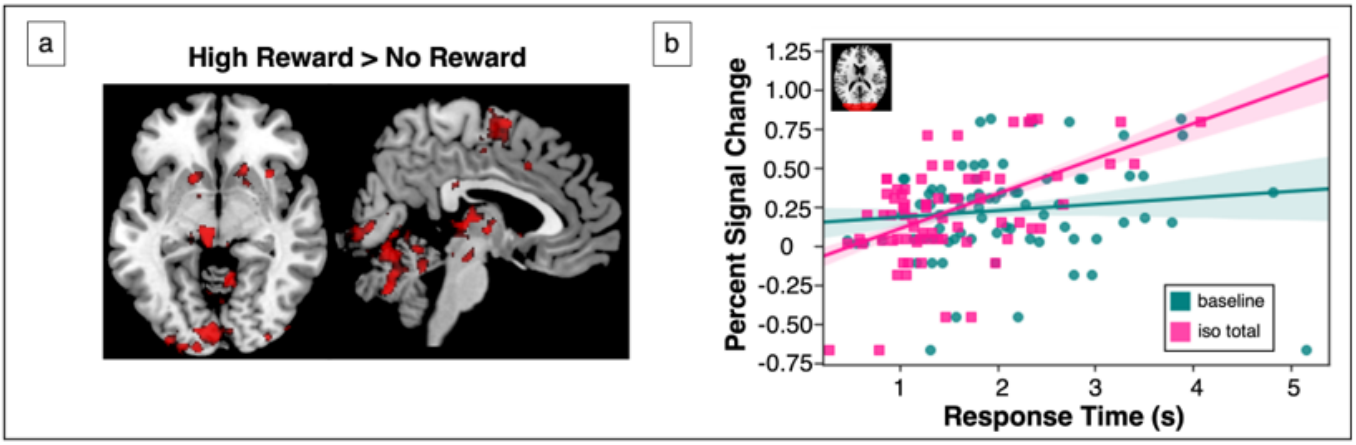

Figure 7

Fig. 8 Neural reward sensitivity predicts effects of isolation on reward seeking.

(a) Univariate group level activity family-wise error (FWE) corrected at the voxel level ($p<0.05$) for the contrast High Reward > No Reward during the anticipation phase of the MID task. (b) Brain reward signals predicting reward seeking in the EBDM after isolation (iso total) but not at baseline for univariate neural reward signals in occipital cortex. Plots depict individual data points, regression lines and confidence intervals (standard errors of the estimate) between neural reward signals and response time in the EBDM task in the baseline (blue) and iso total (pink) session.

Neural reward sensitivity (univariate or multivariate) did not predict reward learning in either the iso total session or the iso media session for any of the ROIs for either social or non-social feedback (all p-values > 0.03).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- NHBSupplementaryMaterialsTomovaetal.March212023.docx