Reviewed Preprint
v1 • August 4, 2026
Not revised

For correspondence:
phonamnp@tcd.ie
gillancl@tcd.ie

Funding: See page 41

Reviewing editor: Nathan Faivre, Centre National de la Recherche Scientifique, France

Dense repeated sampling reveals impaired hierarchical integration of metacognitive confidence that scales with depression

Paveen Phon-Amnuaisuk[1,3], **Vanessa Teckentrup**[1,3], **Sol Fittipaldi**[4,5], **Robert J Crossen**[6], **Celine A Fox**[7], **Anna K Hanlon**[1,2], **Sucharit Katyal**[8], **Stephen M Fleming**[9,10], **Claire M Gillan**[1,4]

[1]School of Psychology, Trinity College Dublin, Dublin, Ireland • [2]Trinity College Institute of Neuroscience, Trinity College Dublin, Dublin, Ireland • [3]FutureNeuro Research Ireland Centre, Trinity College Dublin, Dublin, Ireland • [4]Global Brain Health Institute, Trinity College Dublin, Dublin, Ireland • [5]Cognitive Neuroscience Center, Department of Life and Behavioral Sciences, Universidad de San Andrés, Buenos Aires, Argentina • [6]Department of Engineering, University of Cambridge, Cambridge, United Kingdom • [7]Postdoctoral Research Fellow at the Economic and Social Research Institute, Dublin, Ireland • [8]Department of Psychology, University of Copenhagen, Copenhagen, Denmark • [9]Department of Experimental Psychology and Institute of Cognitive Neuroscience, University College London, London, United Kingdom • [10]Max Planck Centre for Computational Psychiatry and Ageing Research, University College London, London, United Kingdom

eLife Assessment

This **valuable** study uses repeated experience sampling to track the relationship between depressive symptoms and perceptual confidence. The central claim that depression impairs the integration of momentary confidence into a global sense of confidence is supported by **solid**, robust evidence. However, the largely subclinical, highly compliant sample with limited mood variability restricts the generalisability of the conclusions. This study will be of interest to researchers studying mood disorders and metacognition, as it suggests more ecologically valid ways to investigate how mental health influences metacognition.

https://doi.org/10.7554/eLife.112173.1.sa3

Abstract

Depression is consistently associated with metacognitive under-confidence in cross-sectional data. However, the temporal dynamics underlying under-confidence in depression remains unclear. We used ecological momentary assessments (EMA) to examine these dynamics and test candidate mechanisms underlying this relationship. Using a smartphone app, users completed twice-daily EMA assessments of depressive symptoms, and bi-daily assessments of a gamified metacognitive task indexing local (trial-wise) and global (block-wise) confidence. Multi-level vector autoregressive models across 162 citizen scientists (M_{age}=54.4 ± 11.3; 75% female) revealed higher overall between-person depression scores were associated with lower local and global confidence levels – an effect consistently observed at the individual symptom level. Within-person, depression scores significantly covaried with global, but not local, confidence, and cross-lagged analyses found no evidence for temporal precedence in either direction. Turning to mechanisms that might underlie under-confidence in depression, we found overall depression and symptoms to be associated with a diminished integration of local confidence (t-1) into future (t) global confidence (β=-0.280, 90% CrI=[-0.465, -0.096], BF_{10}=136). Moderated mediation analyses with depression at ±1 SD showed that this was mediated by a blunted autoregression of local-confidence (ΔIndex of

Moderated Mediation (IMM)=-0.310, 90% CrI=[-0.533, -0.090], FF_{10}=95). Computational modelling further revealed a more conservative confidence criterion in individuals with higher depression and its symptoms. Together, our results present a novel mechanism for persistent under-confidence in depression: individuals with higher trait depression require more evidence to report high local confidence, while having less stable local confidence fluctuations, resulting in weaker accumulation and integration into future global confidence.

Introduction

Metacognition refers to the ability to reflect on one's own cognition and behaviour (Fleming, 2024) and encompasses constructs such as metacognitive sensitivity (recognising whether individual decisions and actions are correct) and metacognitive bias (overall confidence in performance). These facets of metacognition are crucial for correcting mistakes, improving performance, and deciding whether to engage in activities (Bénon et al., 2024; Fleming & Dolan, 2012). Distortions in metacognition are thought to play a role in several mental health conditions (Hoven et al., 2019), with under-confidence emerging as a central characteristic of depression. Under-confidence has consistently been reported in case-control comparisons of depressive disorders (Hoven et al., 2019) and those scoring higher on an anxious-depression dimension in the general population (Benwell et al., 2022; Hoven et al., 2023; Katyal et al., 2025; Rouault et al., 2018; Seow et al., 2025; Seow & Gillan, 2020).

Recent longitudinal studies have extended these cross-sectional findings showing that depression symptoms reduce over time as confidence increases, both following therapy (Fox et al., 2023) or naturalistically after a two-year follow-up (Agnoli et al., 2023). These findings suggest that lower confidence may be less of a trait vulnerability marker and is instead more of a state-dependent feature of depression, but do not expose how this relationship unfolds over time. Understanding these linkages may have important implications for mechanistic accounts of depression and for developing new avenues for treatment. For example, if low confidence temporally precedes periods of depression, then interventions that address metacognition might be useful for disrupting the cycle of depression. Beyond the temporal dynamics of this relationship, there are also gaps in our understanding of the mechanisms through which confidence in day-to-day decisions contribute to, or are influenced by, broader concepts like self-esteem.

A promising theoretical framework posits that metacognition is structured as a hierarchy. At the lowest level sits our 'local' confidence in specific decisions (e.g. individual trials of a cognitive task), which in turn inform 'global' self-performance estimates (SPEs; e.g. confidence in our overall ability on a task). At even more global levels sit more general beliefs about one's ability or worth as a person (Seow et al., 2021). Recent cross-sectional studies have found that depression is linked more strongly to global measures of confidence and self-beliefs (Hoven et al., 2023; Seow et al., 2025), and that reduced global confidence might act as a prior that negatively biases local confidence (Van Marcke et al., 2024). These findings are consistent with theoretical accounts of reciprocal interactions between levels of a metacognitive hierarchy (Rouault et al., 2019; Seow et al., 2021). Another possibility, however, is that depression may distort the way that local confidence is integrated into global self-beliefs. For example, negative expectations may downweigh positive information when updating higher-order beliefs (i.e., cognitive immunisation; Kube et al., 2020), resulting in persistent negative beliefs, a core feature of depression (Beck, 1979; Kube & Rauch, 2025). Recent experimental work corroborates this view, showing weaker integration of high local to global confidence during a single task session in those with higher levels of anxious-depression (Katyal et al., 2025).

The present study aimed to extend this work and test these ideas using a microlongitudinal design with ecological momentary assessments (EMA) (Myin-Germeys et al., 2018) of metacognition and mental health spaced over 8 weeks. Specifically, we gathered momentary depression ratings twice daily, and metacognitive performance on a gamified perceptual (dot difference) decision making task (Fox et al., 2024) once every second day, via smartphone from N=162 participants. Our primary goal was to isolate within and between-person relationships between depression, local and global confidence, and test for temporal precedence or 'Granger causality' (Epskamp et al., 2018). Thus far, only one study has utilised EMA to investigate confidence and mood (da Fonseca et al., 2023). Using an implicit confidence measure (decision to opt out of a trial), they did not find evidence of cofluctuations between confidence and mood, either contemporaneously or temporally. However, the nonsignificant findings may be due to the small study size conducted over just 10 days, the lack of a subjective confidence measure, and or control for objective accuracy. Interestingly, they reported that metacognition and mood fluctuated over slightly different timescales, on the order of 2.5 and 1 days, respectively. Therefore, a study with a longer overall duration and more widely spaced assessments may be necessary to capture cofluctuations. A secondary aim of the study was to test for any alterations in metacognitive sensitivity in depression. In contrast to confidence biases, studies examining this have typically yielded mixed results, with some studies finding no difference between patients with depression and healthy participants (Agnoli et al., 2023; Culot et al., 2022), while others show better (Rouault et al., 2018) or reduced (Katyal & Fleming, 2026) metacognitive sensitivity along a depression continuum in a healthy sample. We now know that estimating metacognitive sensitivity is more difficult than metacognitive bias and requires hundreds of trials to achieve reliable estimates (Guggenmos, 2021). We therefore sought to leverage our dense repeated samples of metacognition to address this as a secondary study aim; that is, combining data from repeat plays of the task over 8 weeks to achieve the required power to perform robust computational modelling of metacognitive sensitivity and test for associations with depression.

In the present study, we thus investigated i) the temporal dynamics of metacognitive confidence (local and global) and depressive mood; ii) the influence depression has on the integration of local to global confidence; and iii) whether these associations differ at the state and trait level or depending on specific depression symptoms. Extending previous findings, we hypothesised that contemporaneous within-person (i.e. state) and between-person (i.e. trait) depression would be negatively associated with confidence. We further hypothesised that temporal relationships (suggestive of causation) would also exist but did not have a strong *a priori* hypothesis about directionality. Assessing both local and global confidence allowed us to investigate whether depression influences the integration of local to global confidence at the state and trait level across time. Based on Katyal et al. (2025) findings, we hypothesised that depression is linked to poorer integration of local to global confidence. Finally, we leveraged our large volume of per-participant datapoints to fit a computational model to the entire 8-week time series to estimate metacognitive efficiency and biases in post-decisional processing biases in relation to depression (Desender et al., 2022; Katyal & Fleming, 2026).

Results

Depression is associated with lower local and global confidence between-person

We analysed data from 162 citizen scientists (75% female, M_{age} = 54.4 ± 11.3 [19 – 82]) who tracked symptoms of depression twice per day (Fig. 1a) and played a perceptual decision-making task every 2 days over an 8-week period (Fig. 1b). Total depression score (i.e. average of all EMA symptom items) was associated with lower local (Partial correlations (Pcorr) = -0.239, Ps < .05) and global (Pcorr = -0.262, Ps < .05) confidence.

To examine which individual symptoms might drive this effect, we repeated this analysis for all 14 items, correcting for multiple comparisons using the adaptive Benjamini-Hochberg correction (PBH). There was no clear driver of the mean effect, all individual symptoms (except for tiredness) were significantly associated with lower local (-0.263 < Pcorr < -0.170 Ps_{BH} < .05) and global (-0.294 < Pcorr < -0.2, Ps_{BH} < .05) confidence (Fig. 1c). See Supplementary Table 1 and 2 for details. Descriptively, trouble concentrating, sadness and worry had the strongest relationships with both local and global confidence, while feeling guilty, restless or tired consistently showed the weakest association.

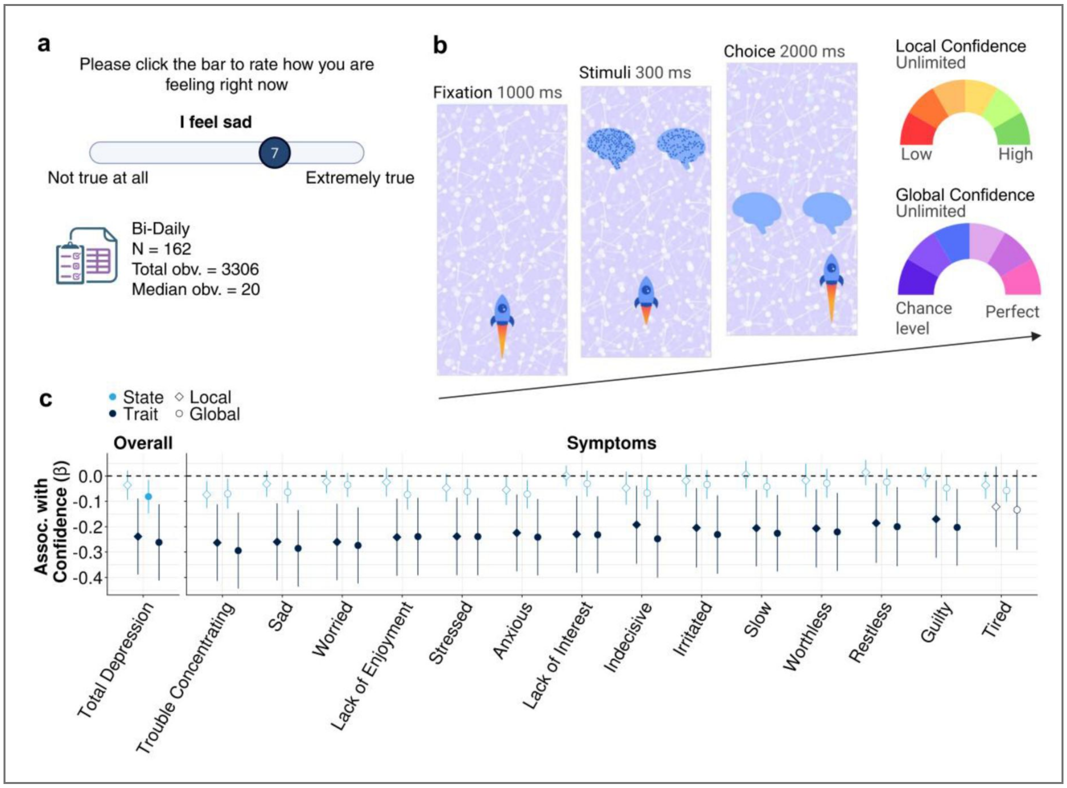


Fig 1. Microlongitudinal design and undirected associations between depression and metacognitive confidence.

(a) Design of the response scale for EMA item and summary of observations used for analysis. 14 depressive symptoms were assessed twice daily (morning and evening) on an 11-point visual analogue scale, while metacognitive confidence was assessed bidaily in the evenings for 8 weeks. (b) Meta-mind procedure. For each assessment, players are instructed to navigate the ship to the side that has more dots. At the start of each meta-mind round (2 rounds with 20 trials each) players are shown a screen with no stimuli and then two icons with varying dot densities then appear at the top of the screen. After the dots disappear, participants move the spaceship and select the icon containing more dots. At the end of each trial, players are asked to rate their local confidence in the accuracy of their choice. After 20 trials, players are asked to evaluate their global confidence in the overall accuracy for the round. (c) Between-person and contemporaneous associations. Plots display results from multilevel autoregressive models examining the relationship between overall depression and its symptoms, with metacognitive local (diamond) and global (circle) confidence. Between-person partial correlation coefficients (in navy, 'Trait') indicate the association between the mean of each EMA item and the mean of confidence, controlling for temporal (autoregressive and cross-lagged) effects as well as age, gender, and education. Within-person contemporaneous partial correlation coefficients (in blue, 'State') indicate how well each EMA item relates to metacognitive confidence in the same time window. Vertical lines reflect the averaged 95% confidence interval. Post hoc analyses of individual symptom level associations were corrected for multiple comparisons using the Benjamini-Hochberg adaptive linear step-up procedure. Filled shapes indicate significant results using an alpha level of $P < .05$.

Older age was linked to lower trait total depression (β = -0.177, SE = 0.07, t = -2.515, p = .013), and lower symptom severity for 11 out of 14 individual depressive symptoms (-0.205 < P < -0. 130, Ps_{BH} < .05). Age was also associated with greater average local (β = 0.166, SE = 0.072, t = 2.308, p = .022) and global (β = 0.204, SE = 0.064, t = 3.179, p = .002) confidence. We found no effects of gender or education (Ps_{BH} > .05).

Within-person fluctuations in depressive mood are associated with changes in global, not local, confidence

We found a significant contemporaneous association between depression and global confidence (Pcorr = -0.081, ps<.023), such that when a person's mood drops, so does their global confidence in the task. This was not the case for local confidence (Pcorr = -0.035, ps>.115). A follow up analysis of individual depression symptoms did not reveal any significant effects, after correcting for multiple comparisons (Ps_{BH} > .05). See Fig. 1c and Supplementary Table 3 and 4 for more details. We found significant autoregressive effects for local (β = 0.278, SE = 0.029, t = 9.352, p < .001) and global (β = 0.098, SE = 0.024, t = 4.001, p < .001) confidence, and to a lesser extent, total depression (β = 0.049, SE = 0.018, t = 2.347, p = .018). Individual depression symptoms also showed significant autoregressive effects (0.11 < p < 0.236, Ps_{BH} < .001). Turning to the question of Granger causality, crosslagged analyses did not reveal any significant relationships between depression and local or global confidence. Follow-up analyses of individual symptoms showed that increases in stress significantly predicted future decreases in global confidence (β = -0.063, SE = 0.022, t = -2.960, PBH = .041). We found no other temporal associations between depression and confidence in either direction (Ps_{BH} > .05; Supplementary Fig 1, Fig 2, Supplementary Table 5 - 8).

Trait depression impairs the integration of local to global confidence

Same-level interactions

Having established that global confidence fluctuates with depression within-person, next we investigated more complex candidate dynamics between mood and different hierarchical levels of confidence that might shed further light on this mechanism. One possibility is that depressed mood diminishes the extent to which local confidence updates or informs global confidence. We tested this by examining the interaction between momentary depression and local confidence in models predicting global confidence at the same time point. We carried out Bayesian analysis as frequentist models failed to converge when using random slopes.

We did not find a moderating effect of momentary depression on the extent to which local confidence related to global confidence contemporaneously (β = 0.268, 90% CrI = [-0.071, 0.607], BF_{10} = 1/9). In an exploratory analysis of individual symptoms, we did however see a contemporaneous interaction effect with worthlessness (β = 0.416, 90% CrI = [0.138, 0.698], BF_{10} = 1/161), feeling guilty (β = 0.301, 90% CrI = [0.036, 0.565], BF_{10} = 1/32), and trouble concentrating (β = 0.463, 90% CrI = [0.231, 0.699], BF_{10} = 1/1141). Contrary to our hypothesis, these factors increased the strength of the association between local and global confidence. Similarly, there was anecdotal to no evidence supporting the moderating effect of trait depression and its symptoms on the association between local and global confidence cross-sectionally (β = 0.024, 90% CrI = [-0.106, 0.155], BF_{10} = 1; -0.058 < p < 0.132, BF_{10} < 1/3). See Supplementary Table 9 and 10 for details.

We next explored how these dynamics unfold over time, by testing how momentary depression impacts the forward propagation of local confidence to global confidence at the next time point. We did not find any evidence for an effect of depression (β = 0.185, 90% CrI = [-0.226, 0.599], BF_{10} = 1/3). However momentary increases in feelings of indecisiveness (β = 0.463, 90% CrI = [0.136, 0.795], BF_{10} = 1/89), guilt (β = 0.305, 90% CrI = [0.003, 0.609], BF_{10} = 1/20), and worthlessness (β = 0.443, 90% CrI = [0.089, 0.787], BF_{10} = 1/49) specifically promoted the integration of local to future global confidence. See Supplementary Table 11 for details.

Cross-level interactions

The previous results concerned momentary fluctuations in depressed mood from one's own set point. Next, we examined how *trait* depression relates to how local confidence integrates into global confidence. We did not find evidence for contemporaneous associations (Supplementary Table 12 for details), but for lagged analyses, we found significant and extreme evidence that trait total depression (β = -0.279, 90% CrI = [-0.467, -0.093], BF_{10} = 156) was associated with *reduced* integration of local confidence into global confidence at the next time point. This finding was consistent across all depressive symptoms (-0.366 < β < -0.199, BF_{10} > 20). See Fig. 2a and Supplementary Table 13 for details.

One possible mechanism for this may be that trait depression makes experiences of local confidence less sustained, i.e. local confidence is more quickly pulled down towards a set point of trait under-confidence, thus weakening subsequent integration of local into global confidence. Indeed, a two-way interaction found depression was associated with a significantly weaker autoregression of local confidence (β = -0.072, 90% CrI = [-0.123, - 0.022], BF_{10} = 97). With the exception of tiredness and slowness, this effect was strong and consistent across depressive symptoms (-0.092 < β < -0.057, BF_{10} > 23). We additionally modelled residual within-person variance in local confidence as a function of depression while including a participant-level random intercept. This tested whether depression was associated with greater within-person variability in local confidence, above and beyond the variability at the participant level. However, this model resulted in similar LOO-CV performance to a random-intercept-only model, suggesting depression did not meaningfully explain variability in within-person local confidence. See Supplementary Table 14 and Note 2 for details.

Moderated mediation analysis

To further dissect the mechanism underlying this effect, we conducted a moderated mediation analysis (Hayes, 2015). See Fig. 2b for details. We tested if the significant cross-level interaction between trait depression and lagged local confidence (Path c), could explained by an indirect pathway where depression moderated the autocorrelation of local confidence (Path a) together with a cascading effect on the integration of contemporaneous local to global confidence (Path b), while controlling for lagged depression (Path c').

To test this, we calculated the difference in the index of moderated mediation (ΔIMM) for depression at ±1 SD (i.e., difference in indirect effects at different levels of depression), and hypothesised a negative ΔIMM (reflecting poorer integration that scales with depression) across total depression and its symptoms. The results indicated moderate to extreme evidence in support of the moderated mediation process for total depression (ΔIMM = -0.310, 90% CrI = [-0.533, -0.090], BF_{10} = 95) and individual symptoms (-0.393 < ΔIMM < -0.272, BF_{10} > 20), except for slowness and tiredness. See Fig. 2c and Supplementary Table 15 for details. This appeared as a complete mediation for total depression and most symptoms except for restless (β = -0.201, 90% CrI = [-0.346, -0.059], BF_{10} = 100) and stressed (β = -0.15, 90% CrI = [-0.294, -0.011], BF_{10} = 25) which showed a strong evidence of partial mediation process through Path c'. See Supplementary Table 16 for details.

Overall, the findings suggest that trait depression blunts the autocorrelation of local confidence, which subsequently impairs how local confidence is integrated into global confidence. See Fig 2b for details on the process. This is consistent with a mechanistic explanation for how negative beliefs in depression are formed and maintained over time.

To interrogate this further, we performed a three-way interaction between trait depression, lagged local confidence, and valence (i.e. whether confidence was above or below a person's mean). We found moderate but insufficient evidence supporting a weaker carry-over effect of local confidence when it fluctuated above the mean in total depression (β = -0.016, 90% CrI = [-0.046, 0.015], BF_{10} = 4), and this effect was similar across most symptoms (-0.009 < β < 0.006, BF_{10} < 11). This suggests that for total depression and most symptoms, the impairment is not dependent on the valence of local confidence fluctuations and is best described as a *regression to the mean* effect. However, there was very strong evidence for an interaction in models of slowness (β = -0.045, 90%

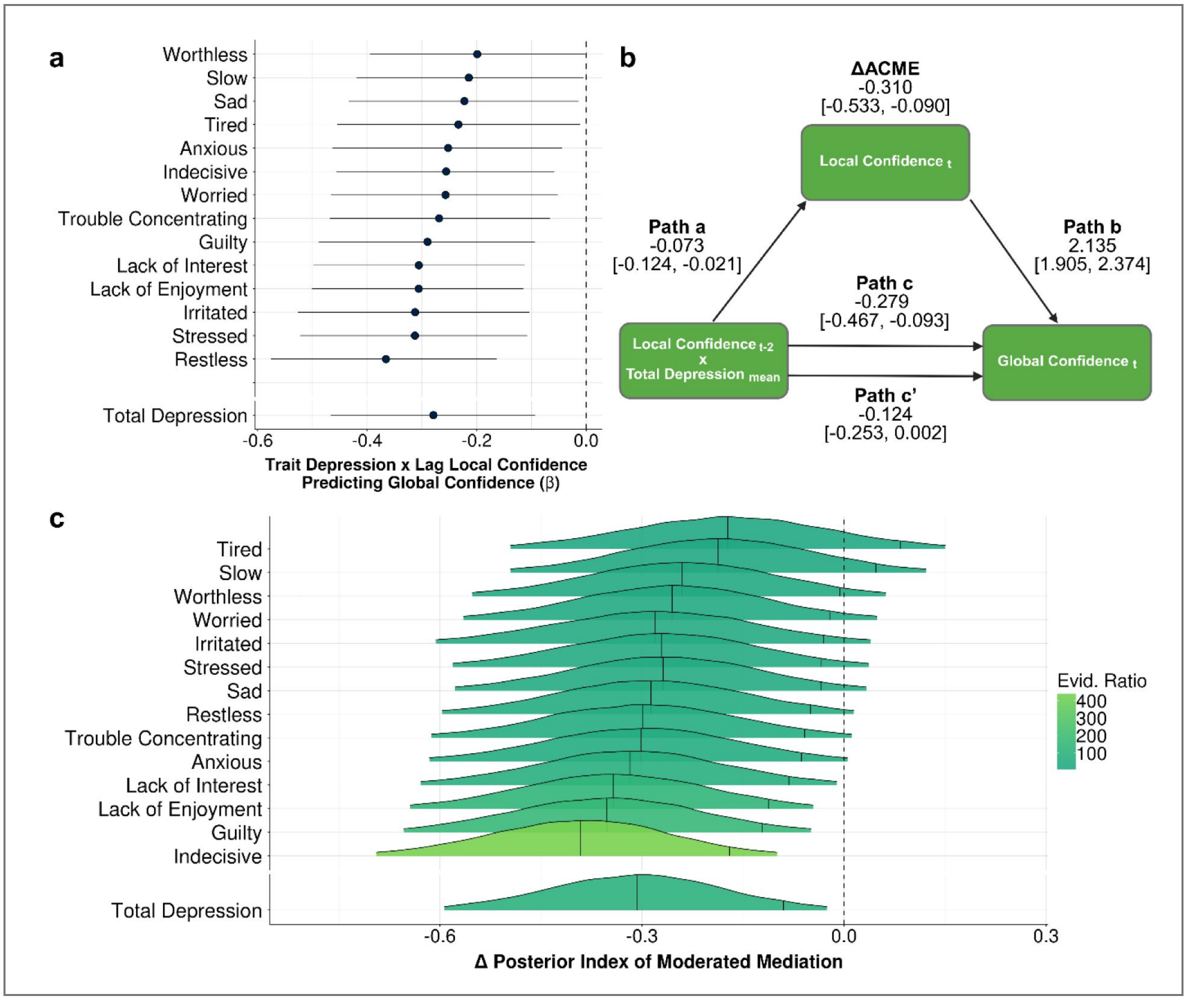


Fig. 2. Trait depression weakens the carry-over from local to global confidence, and this operates via a moderated-mediation pathway.

(a) Cross-level interaction between trait depression/symptoms (between-person) and lagged local confidence (within-person deviation) predicting global confidence. Point estimates of standardised beta interaction coefficients from Bayesian mixed-effects models for total depression and each symptom, with horizontal lines indicating the 90% credible interval. Filled shapes indicate 90% credible intervals are outside of 0. (b) Moderated mediation path diagram for total depression. Trait total depression interacts with within-person lagged local confidence to predict within-person concurrent local confidence (*Path a*) and in turn global confidence (*Path b*). The same interaction term has a total effect on global confidence (*path c*) and a direct effect controlling for concurrent local confidence (*Path c'*). Values under each path reflects the standardized path coefficients with 90% credible intervals. ΔACME above the mediator denotes the index of moderated mediation, i.e., the difference in the indirect effect (a x b) at ±1 SD of trait total depression. (c) Exploratory moderated-mediation test. Plots show the difference in the posterior distributions of index of moderated mediation (ΔIMM) at ±1 SD of trait depression. Negative ΔIMM values indicate weaker indirect effects at higher depression. Fill indicates the Bayesian evidence ratio (Savage–Dickey Bayes factor), with lighter greens denoting stronger evidence. For total depression and symptoms (except for tiredness), ΔIMM < 0 with strong–extreme Bayesian support, consistent with depression blunting local-confidence autocorrelation and, in turn, its integration into global confidence.

CrI = [-0.078, 0.011], BF_{10} = 70) and trouble concentrating (β = -0.044, 90% CrI = [-0.076, 0.012], BF_{10} = 77), indicating a weaker sustained fluctuation of above-mean local confidence specifically at higher levels of these symptoms. Finally, we tested whether the autoregressive effect of local confidence was valence dependent using a two-way interaction, and only found moderate evidence that autocorrelation was stronger following positive compared to negative within-person fluctuations in local confidence (β = 0.019, 90% CrI = [-0.009, 0.048], BF_{10} = 6), indicating at most, weak asymmetry.

Under-confidence in depression is associated with computational markers signalling negative expectations

Finally, we investigated how aberrations in post-decisional processing in depression might generate distorted local confidence. We first examined metacognitive efficiency using v-ratio, i.e., the ratio of the amount of evidence used to estimate confidence, to that used to make the perceptual decision. A v-ratio of 1 would imply that the evidence used to make perceptual decisions and estimate confidence is equal, i.e. perfect metacognitive efficiency. We also investigated distortions in the accumulation of post-decisional evidence (v-bias), additive distortions in the shift of confidence criterion (a-bias), and multiplicative shifts in the distortions of confidence criterion (m-bias).

We did not find any association between v-ratio and depression (β = -0.041, SE = 0.08, t = - 0.518, p = 0.605) or its symptoms (*PsBH* > .05). Instead, we found that depression was associated with greater a-bias (β = 0.381, SE = 0.120, t = 3.167, p = .002), which corresponds to a more conservative threshold for reporting high confidence (Fig 3). With the exception of feeling guilty or tired, the same relationship with a-bias was observed for all individual symptoms of depression (0.276 < β <0.378, *PsBH* < .05) in the following order of effect size: trouble concentrating, feeling worried, lack of interest, feeling stressed, lack of enjoyment, feeling slow, feeling sad, feeling irritated, feeling indecisive, feeling anxious, feeling restless, feeling worthless. We did not find associations between overall depression and other candidate distortions (v-bias, a negative bias in post-decision drift rate) *or* (m-bias, a slower rate of change from evidence to confidence; *PsBH* > .05). See Fig. 3d for details and Supplementary Table 17 for symptom, parameter, and demographic associations.

Overall, these findings suggest that participants with higher depression require more evidence (higher expectations or more conservative evidence threshold) to achieve the same level of confidence as healthier participants with the same objective performance.

Discussion

Distortions in metacognition are a characteristic feature of several mental health disorders (Hoven et al., 2019). In depression, these distortions manifest as poorer metacognitive beliefs, global under-confidence, and local under-confidence (Benwell et al., 2022; Hoven et al., 2019, 2023; Rouault et al., 2018; Seow et al., 2025; Seow & Gillan, 2020), signaling a pervasive negative bias which manifests across levels of a metacognitive hierarchy. Prior work suggests that depression may not only vary over time with under-confidence (Agnoli et al., 2023; Fox et al., 2023) but also impair the integration of local confidence into global confidence (Katyal et al., 2025). However, the mechanistic processes underlying these relationships remain unclear as they have thus far only been assessed within the context of a single task. Here, we used a microlongitudinal design to ask i) whether within-person fluctuations in depressive mood precede under-confidence or vice versa, and ii) whether depression symptoms impair the integration of local confidence to global confidence across time. We found that within-person increases in depressive mood were accompanied by underconfidence at the same time point, but neither change preceded the other. Additionally, we found that depression was associated with an increase in confidence threshold, characteristic of defensive pessimism (Norem & Cantor, 1986). Crucially, individuals with higher levels of depression exhibited weaker autocorrelation of local confidence, which cascaded into poorer concurrent integration of local to global confidence. In other words, pervasive global underconfidence in depression may be driven by difficulties

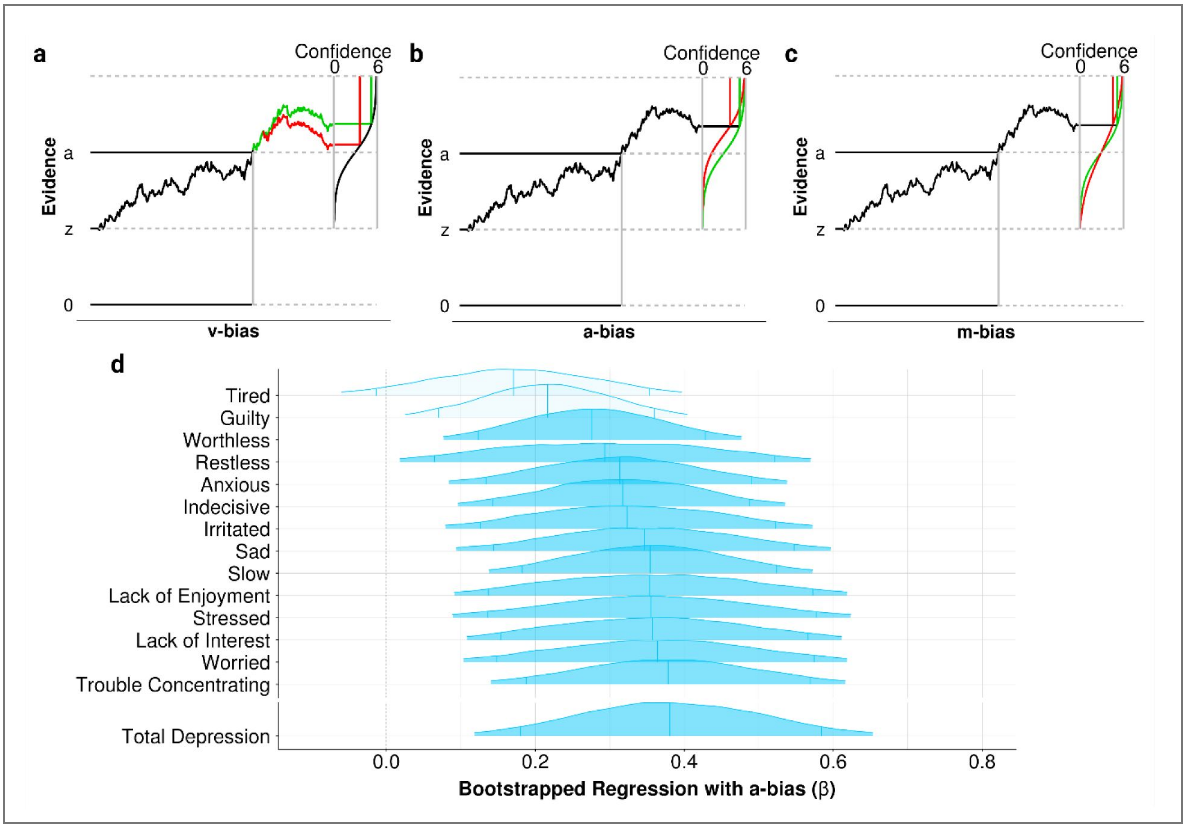


Fig 3. **Computational accounts of distorted post-decision evidence accumulation for the case of under-confidence in relation to depression.**

(a) v-bias: a distorted accumulation of evidence over time. A negative bias in post-decisional drift rate (red) results in lower evidence favouring the decision made, resulting in lower confidence. (b) a-bias: a distorted shift in confidence criterion. An additive shift in the confidence criterion (red) - represented by a lateral shift in the sigmoid function translating evidence into bounded confidence, results in the same amount of evidence translating into lower confidence. (c) m-bias: a multiplicative shift in confidence criterion. A multiplicative shift in the confidence criterion (red) – represented by a flatter sigmoid function, results in a slower rate of change from accumulated evidence into bounded confidence. (d) Bootstrapped mixed effects models predicting depression from a-bias. Ridgelines show bootstrapped standardised beta coefficient distribution for parameter a-bias predicting depression, with $5^{th}/50^{th}/95^{th}$ quantile lines. Shading of the ridges indicates frequentist significance. Analyses included all computational parameters and controlled for age, gender, and education as covariates. Symptom-level associations (but not total depression) were corrected for multiple comparisons.

sustaining high levels of local confidence over time. Overall, our findings highlight how local and global under-confidence are characterised by distinct mechanisms operating at different time scales.

There is causal evidence from experimental inductions of sad (happy) mood that suggest alterations to mood can cause changes in metacognitive under-confidence (overconfidence) (Culot et al., 2021; Koellinger & Treffers, 2015). Here, we sought to extend this work using real world depression fluctuations (rather than mood induction) and tested for an important pre-requisite for causation, temporal precedence. Contrary to our expectations, we found little to no evidence that depression and confidence predicted each other at a two-day lag. Instead, we observed a small within-person covariation between total depression and global underconfidence at the same time point (Pcorr = -0.081). The absence of Granger causality may be explained by a mismatch between the temporal resolution. Put simply, temporal relationships at the scale of days may be too large to reliably establish Granger causality between mood and under-confidence. Supporting this, we found that depressive mood and metacognitive confidence exhibited different autocorrelations, in line with prior work (da Fonseca et al., 2023). Depression showed the weakest carry-over (β = 0.049), followed by global (β = 0.098) and then local confidence (β = 0.278). Therefore, it is likely that any top-down mood-driven changes in confidence decay rapidly as new mood states emerge. Alternatively, day-to-day mood fluctuations in an unselected general population sample may be too small and brief (Thompson et al., 2021) to exert a lasting effect on metacognitive confidence. For this reason, future studies should consider a clinical sample with larger affective variability. Future work should also explore the resolution with which depression and confidence co-fluctuate. More broadly, the significant contemporaneous association between global confidence and total depression, together with similar carry-over magnitude between these constructs, supports the view that depression is more tightly linked to global metacognitive constructs than to local confidence (Hoven et al., 2023; Seow et al., 2025).

Between-person analyses indicate that higher levels of overall depression and most depressive symptoms (except tiredness) were associated with lower average local and global confidence. This pattern is consistent with prior cross-sectional findings in clinical and non-clinical samples (Hoven et al., 2019).

Leveraging the large number of trials gathered per person over the 8 weeks of the study, we applied computational modelling (Katyal & Fleming, 2026) to gain insight into biases in post-decisional processing that may contribute to under-confidence. Higher levels of overall depression and its symptoms (except for tired and guilty) were associated with a more conservative confidence criterion, that is, requiring more evidence to report high confidence. This pattern is consistent with accounts of defensive pessimism, whereby more depressed individuals set higher standards or lower expectations to avoid disappointment (Norem & Cantor, 1986; Villano et al., 2023; Villano & Heller, 2024), and with predictive coding accounts of depression emphasising persistent negative expectation (more rigid negative priors) in depression (Kube et al., 2019, 2020). Consequently, it is plausible that underconfidence in depressed individuals may be addressed by targeting these rigid negative expectations. Nonetheless, because the computational model was fit to each participant's full time series collapsed across days, we could only extract between-person snapshots of post-decisional dynamics. Future studies should investigate how these processes unfold at the within-person level over time.

Past work has established that local confidence integrates into global beliefs (Rouault et al., 2019), and one study suggested that depression may impair this integration during a single testing session (Katyal et al., 2025). Expanding on this finding, we tested whether the integration of local to global confidence would be similarly blunted over time. Using a crosslagged interaction and moderated mediation analysis, we found moderate to extreme evidence that higher trait depression and most symptoms (except worthlessness) were associated with a reduced carry-over of local confidence from the last assessment, subsequently impairing the contemporaneous integration of local into global confidence. This provides a plausible mechanism for how under-confidence and negative self-beliefs are maintained over time in depression. Specifically, a low autocorrelation means that momentary surges in local confidence are unlikely to persist and accumulate into positive global beliefs in those with higher depression. This is consistent with hypotheses of overprecise negative priors in predictive coding accounts of depression (Kube et al., 2020) - whereby such priors may constrain positive within-person fluctuations of local confidence back toward trait underconfidence, manifesting as a "*regression to the mean*" effect. Alternatively, individuals with higher depression may have more volatile internal representations of confidence over time, consistent with accounts of heightened sensitivity to integrate postdecision evidence (Moses-Payne et al., 2019). To test this, we additionally specified within-person residual variance of local confidence as a function of trait total depression in a separate model but did not find a meaningful improvement in the model's predictive performance compared to a random-intercept only model. This suggests a pattern that is consistent with a stronger *"regression to the mean"* effect in high depression.

We then aimed to investigate whether impaired integration disproportionately targets positively valenced information (Kube et al., 2020), as shown in prior work (Hobbs et al., 2022; Katyal et al., 2025; Korn et al., 2014; Villano & Heller, 2024). However, our test for a three-way interaction did not replicate a valence-specific impairment that scales with depression, except in relation to symptoms of slowness and trouble concentrating. The findings suggest that depression may be responsible for a broader impairment to the temporal persistence of local confidence, rather than a valence specific deficit. There was also limited evidence to support stronger carry-over effects for positive fluctuations of within-person local confidence. Although modest, this pattern is broadly consistent with characterizations of optimism-bias in the general population (Lefebvre et al., 2017; Sharot et al., 2011; Sharot & Garrett, 2016; Villano et al., 2023).

We conducted exploratory symptom-level analyses and observed a surprisingly high degree of homogeneity across symptoms. This adds to the current conversation on the validity of the construct of depression and the extent to which individual symptom dynamics ought to be modelled (Fried et al., 2016; Fried & Nesse, 2015; Zimmerman et al., 2015), at least in the context of confidence. In terms of demographics, we also found older age to be associated with lower total depression, and greater local and global confidence. This is in line with previous literature noting improved mental health (Thomas et al., 2016) in older adulthood, but contrasts with other findings of lower confidence with age (Katyal & Fleming, 2026; McWilliams et al., 2023). We did not observe gender differences (women compared to men) in under-confidence either, in contrast to prior work (Hoogervorst et al., 2024; Katyal & Fleming, 2026). These discrepancies may reflect our relatively small, older (M_{age} = 54.4 ± 11.3) and predominantly female (75%) demographic which contrasts the larger, younger, and more balanced gender demographics typical of metacognition research (Xue et al., 2024).

Several limitations should be considered when interpreting these findings. Firstly, limited variability in mood fluctuations in the healthy general population sample may have attenuated within-person associations between under-confidence and depression. This is especially relevant given that the inclusion criteria required participants to be compliant and engaged in an 8-week study. Patient samples will exhibit a greater range and more variation in symptoms and this could reveal associations with metacognition which are distinct from healthier participants. Second, our data can only speak to fluctuations occurring across a two-day window; future research should explore whether the temporal dynamics of depression and confidence can be better characterized across alternative time scales. Third, because metacognitive constructs dynamically interact with each other across hierarchical levels (Rouault et al., 2019; Seow et al., 2021), the absence of high-level metacognitive beliefs (i.e., self-beliefs) in our analyses may, at worst, confound our associations, and at best, provide an incomplete account of the dynamics between metacognition and depression. Finally, our within-person "depression" variable reflects short-term fluctuations in symptom ratings around each individual's mean, whereas "depressive disorders" are characterised by their persistence over time (Fifth Edition, Text Revision, DSM-V-TR). In this sense, our within-person analyses may not accurately reflect or generalize to longer-term trajectories associated with the onset, maintenance, and remission of depression and its symptoms.

In summary, our findings suggest that higher trait-level depression is associated with underconfidence partly because local confidence is less stable over time, and therefore less able to accumulate and integrate into later global confidence. This presents a temporal perspective into the emergence and persistence of global under-confidence in depression. More broadly, within-person fluctuations in mood may be better understood as a contemporaneous correlate of under-confidence rather than a causal component. At the trait level, higher depression was robustly associated with local and global under-confidence, and the former may be explained by a more conservative confidence criterion characteristic of defensive pessimism. Together, these findings suggest that preventative measures aimed at depression may benefit from stabilizing positive local confidence fluctuations, thereby improving integration of local confidence across metacognitive hierarchies.

Methods

Participants

N = 162 participants post-inclusion criteria (75% female; 23% male) aged between 19 and 82 years (M = 54.38 ± 11.34) were recruited via the Neureka app from July 2022 to June 2025 (see Supplementary Table 18 for additional demographic information). We term these participants 'citizen scientists', as they received no compensation for their participation.

EMA assessment and validation

Participants took part in the 'Brain Changer' science challenge within the Neureka app, includes an EMA module tracking symptoms of depression and a perceptual decision-making task to measure metacognitive confidence. After an initial baseline assessment, participants were required to complete an 8-week EMA procedure tracking depressive mood symptoms and metacognitive confidence. At sign up, participants were required to select two 3-hour time windows (between 6:00 - 11:30 AM and 6:00 - 11:30 PM) to receive notifications for their assessment. Notifications were delivered at the start of each time window and participants were allowed to respond until the end of the selected time window. Each notification prompted participants to rate their experience on a sliding scale ranging from 0 (Not true at all) to 10 (Extremely true). The EMA items were adapted from the Quick Inventory of Depressive Symptomatology, self-report (QIDS-SR-16; Rush et al., 2016), a validated questionnaire to evaluate the Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition, Text Revision, DSM-IV-TR) depression criterion symptom domains. The item on suicidality was not included. Three items pertaining to trouble sleeping, excessive sleeping, and appetite were only assessed in the morning and excluded from all analyses, which focused on afternoon assessments that coincided with the cognitive task.

To validate our EMA questionnaire, we correlated the EMA assessed in the morning across the second week with QIDS-SR-16 assessed on assessment day 14 and found it to be excellent (r = 0.71, 95% CI = [0.62, 0.79], $p < .001$), indicating good convergent validity (see Fig. 4a). Overall, depression scores were stable across each participant's first 14 evenings (ICC = 0.888, 95% CI = [0.864, 0.910], $p < .001$; see Fig 4b and c), though this was perhaps exaggerated by low variance as depression was zero-inflated across participants throughout the study (see Fig. 4b). There was no systematic trend towards improvement or worsening over the 8 weeks of the study (β = -0.019, SE = 0.013, t = -1.416, p = .159).

Metacognitive confidence was assessed bi-daily in the evenings using a game in the app called 'Meta-mind' - a gamified version of a visuo-perceptual metacognitive task. 'Metamind' has been described and validated in a previous work (Fox et al., 2024). In this game, players are presented with two moving icons (left and right side) that differ in the number of dots contained within them and descend the screen. Players are instructed to navigate a spaceship to the icon containing more dots by touching the screen and allowing the ship to collide with the icon. After a choice is made, players are asked to rate their local confidence in the accuracy of their choice on a 1-6 scale ranging from 'Low' to 'High'. Reaction times were measured starting from the appearance of the

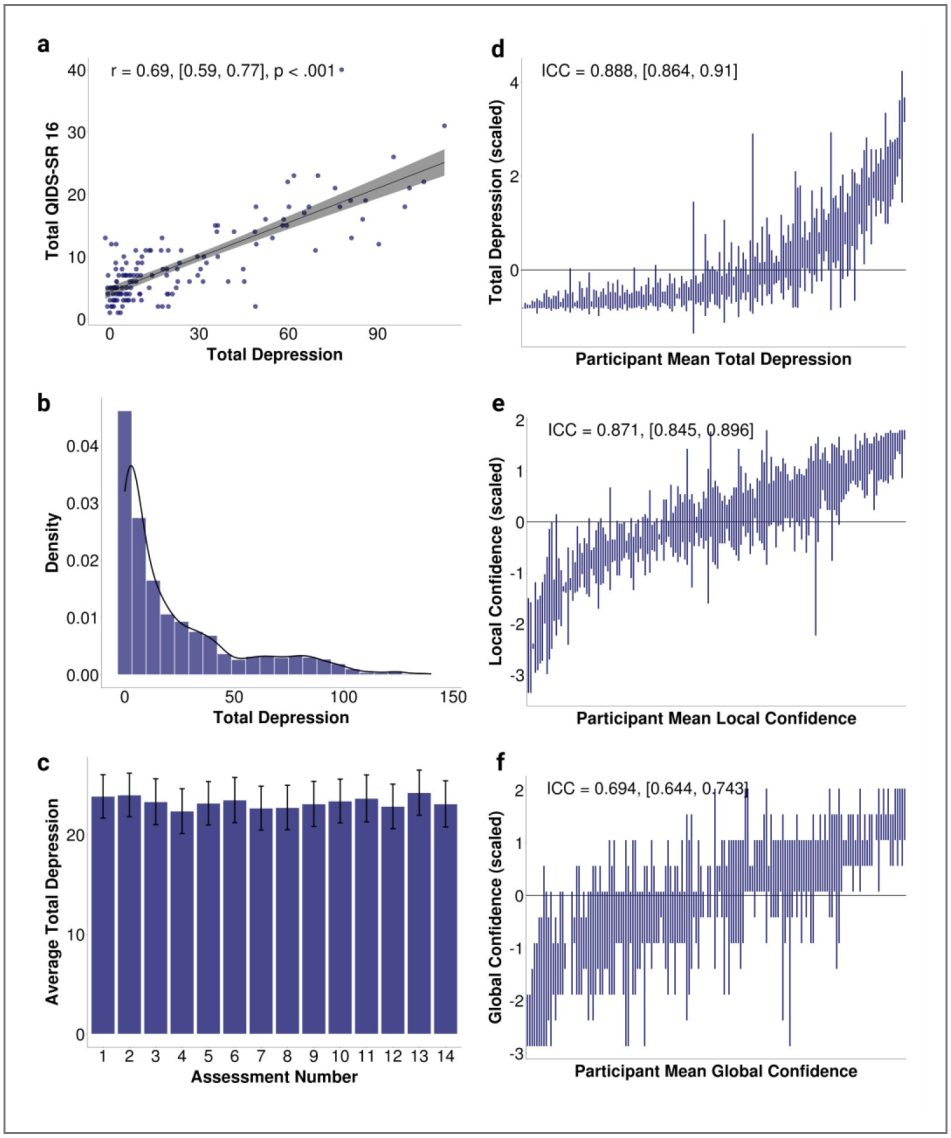


Fig 4. Reliability and validity checks for depression EMA and confidence ratings.

(a) Bivariate correlations. Between-person Spearman's rank correlation coefficients between non-detrended total depression scores in the morning (second week) and total QIDS-SR-16 scores assessed on assessment day 14. (b) Distribution of the sum of all depression EMA scores (i.e. total depression) across each participant's first 14 evening assessment. Total depression scores show zero-inflation. (c) Average total depression over each participant's first 14 evening assessments. There is minimal drift in average total depression scores in 162 participants across assessments. The error bars represent standard error. Internal consistency of total depression (d), local confidence (e), and global confidence (f) ratings for each participant's first 14 assessments ordered by their means. ICC analyses were performed using evening EMA assessments of 162 participants with a total of 2268 observations.

prompt, up until when the final confidence decision was made. After 20 trials, players evaluate their overall accuracy for that round on a 1-6 scale ranging from ‘Chance Level’ to ‘Perfect’. This block-by-block measure indexed global confidence. No feedback is provided to participants. The task uses a ‘two-down-one-up-staircase which maintains the accuracy within a range of 0.6 to 0.85. The initial staircasing was performed during a baseline testing session across 100 trials which has 20 burn-in trials before the staircase ‘settled’ within the desired accuracy range. The task difficulty level for each participant at the end of a session was always carried over to the next testing session to avoid restarting the staircase (and needing to discard more burn in trials). After the baseline assessment, participants completed 40 trials per day, and the staircase continue to operate throughout to account for any gradual improvements in objective performance that would confound confidence estimates.

To validate the staircasing procedure over the time series we fit a model comparing the accuracy from the 1st (assessment after baseline) and 15th assessment day and found no significant improvement in accuracy (β = 0.137, SE = 0.12, t = 1.144, p = .255). Furthermore, we did not find any significant associations between time and local (β = 0.013, SE = 0.016, t = 0.776, p = .439) or global confidence (β = -0.008, SE = 0.015, t = -0.497, p = .62). However, as expected, difficulty of the task increased over time (decreased log dot difference) as participants became more proficient (β = -0.042, SE = 0.02, t = -2.047, p = .042). The stability of the confidence data was very high for local confidence (Fig 4d; ICC = 0.871, 95% CI = [0.841, 0.896], p < .001), but only fair for global confidence (Fig 4e; ICC = 0.694, 95% CI = [0.644, 0.743], p < .001), although this can be attributed to the latter’s fewer datapoints (2 vs 40 per assessment). Overall, this suggests that the staircase was successful and this change in objective perceptual decision-making performance did not translate into changes in accuracy or confidence.

Data preparation

Behavioural outcomes and exclusions

N = 976 participants with a total of 6784 assessments (range = 1 – 28 assessments per person) were included for initial pre-processing. We began by excluding individual assessments (i.e. a single day of task data) where participants had selected the left or right option more than 95% of the time; 5 assessments across 5 participants (1 assessment each) failed this criterion across the study duration, resulting in 1 participant being excluded. Then, assessments that achieved accuracy rates < 60% or > 85% were excluded, suggesting the staircase had failed to converge for them; 182 assessments across 148 participants (range = 1 – 19 assessments per person) met this exclusion criterion, resulting in 11 participants being removed. Individual trials in an assessment were also excluded if participant’s choice or confidence reaction times were implausibly fast, i.e. < 150ms (81 trials across 33 participants, range = 1 – 18 trials per person) or slow, i.e. > 5s (560 trials across 204 participants, range = 1 – 26 trials per person). Following that, assessments with < 40 trials or 2 blocks of assessments for that day were also removed. This was true for 6597 assessments across 964 participants (range = 1 - 28 assessments per person), resulting in the removal of 55 participants. Following the trial and day level exclusions, participant-level data was excluded if their confidence reports had less than 0.17 standard deviation (1/confidence levels) across the whole time series (18 participants removed), and if they did not have at least 50% (14/28) of the ‘Meta-mind’ assessments and the corresponding EMA items at the same time point (727 participants removed). An additional 2 participants were removed due to missing demographic information.

After exclusions, we had a total of 7408 (6906) evening (morning) EMA assessments across 162 participants, of which 3306 coincided with metacognition assessments (median assessment per person = 20, range = 14 – 28). Mean bi-daily evening EMA responses and confidence ratings are provided in Supplementary Table 19, and data distribution are displayed in Supplementary Figure 1.

Detrending

The Kwiatkowski-Phillips-Schmidt-Shin (KPSS; Kwiatkowski et al., 1992) showed that on average, 75% (71% - 78%) of depression EMA items were stationary across participants. In contrast, 100% of local and global confidence data were stationary. However, total depression (sum of depression EMA items) was only stationary across 66% of participants. Given the accepted threshold of 70% (Bringmann et al., 2013), total depression was detrended to ensure stationarity. We regressed each participant's total depression rating against a linear and exponential term (-0.2) of 'assessment day' (i.e., the number of days since baseline of the study). The residuals from this model were added to each participant's mean score across the entire assessment period, and these detrended time series were then used in further analyses. Post detrending, KPSS showed that 95% of the depression items were stationary.

Statistical analysis

All participants provided informed consent in accordance with the European General Data Protection Regulation (GDPR). Data were pre-processed and analysed using R Statistical Software (v4.4.0) contained within a Docker container (Boettiger & Eddelbuettel, 2017; Nüst, Eddelbuettel, et al., 2020; Nüst, Sochat, et al., 2020).

Multilevel vector autoregressive models

We used a two-step multilevel vector autoregressive model using the 'lme4' package (Bates et al., 2015) to investigate within (temporal and contemporaneous) and between-person associations between depression and metacognitive confidence. Analysis was performed for local and global measures of metacognitive confidence, total depression (sum of all EMA items per day) and for each depression item. We adapted the methodology employed by Epskamp et al. (2018) to incorporate covariates as the mlVAR package does not allow for their inclusion. All models included a random intercept to account for individual differences in the dependent variable. Random slopes were not included in any of the analysis as the models failed to converge. Confidence and depression variables were scaled across the sample, and predictors were also within-person centred. Age (continuous in years), gender (factor; baseline = male), and education (binary: below-university level vs university) were entered as covariates and scaled. Below-university includes lower and upper secondary education, and no formal education, while university includes bachelor's, master's, and PhD degrees or equivalents. Gender and education were coded using a custom contrast such that the intercept would reflect the grand mean, while non-baseline levels were compared to variable's respective baseline (i.e., female and no university level education).

Separate analyses were conducted for local and global confidence. We focused primarily on depression scores, i.e. the sum of all individual EMA items, but complemented those analyses with exploratory analyses of individual symptoms. The generic models used for these analyses were as follows (1,2):

$$\begin{aligned} Confidence_t \sim Depression_{t-2} + Confidence_{t-2} + Depression\_mean + Age \\ + Gender + Education + (1|PlayerID) \end{aligned} \tag{1}$$

$$\begin{aligned} Depression_t \sim Confidence_{t-2} + Depression_{t-2} + Confidence\_mean + Age \\ + Gender + Education + (1|PlayerID) \end{aligned} \tag{2}$$

Because metacognition was assessed bi-daily, the temporal associations were estimated at a two-day lag (t-2). Between-person partial correlations were estimated by averaging the regression coefficient (Depression_mean/Confidence_mean) in the forward and reverse analyses (Epskamp et al., 2018). Contemporaneous within-person associations were estimated by correlating the residuals of the forward regression and reverse regression. Because the temporal and trait associations have been partialed out, the residuals are assumed to represent a contemporaneous link (Epskamp et al., 2018). The models used for analyses is as follows:

$$Forward\_residuals \sim 0 + Reverse\_residuals + (0 + Reverse\_residuals|PlayerID) \tag{3}$$

$$Reverse\_residuals \sim 0 + Forward\_residuals + (0 + Forward\_residuals|PlayerID) \tag{4}$$

P values for the exploratory symptom-level analyses were corrected for multiple comparisons using the Benjamini-Hochberg adaptive linear step-up procedure (Benjamini et al., 2006) with an alpha level of $P < .05$. This approach aims to reduce type-1 errors (false positives) whilst accounting for the positive correlations in the test-statistics of the outcomes, e.g. symptom level analyses where associations are expected to be similar to each other. Using the "AND" rule, between and within-person contemporaneous associations were only considered significant if both regression parameters used for averaging are $P_{BH} < .05$. Bayesian models were not used to address the lack of random slopes in the chosen mlVAR framework as the approach required the ordinal outcomes to be scaled (Epskamp et al., 2018), which in turn, required the selection of likelihood families inconsistent with the data generating process.

Bayesian interaction models

We additionally ran Bayesian ordinal regressions for the interaction analyses using the brms package (Bürkner, 2017). This allowed us to fit random slopes which failed to converge using frequentist methods, while also allowing us to model the data generation process more accurately (Bürkner & Vuorre, 2019; Liddell & Kruschke, 2018). Unlike the frequentist models, global confidence as the dependent variable was not scaled. The Bayesian workflow recommended by Gelman et al. (2020) was used to validate the computations. With the exception of model 12, we could not conduct cross validation analyses due to compute limitations and kept the models maximal. Details on the weakly informative priors used can be found in Supplementary Note 1.

A cumulative probit family was used for the ordinal regression analysis with 8 chains, 2000 warmup samples and 6000 post warmup iterations. Divergent transitions were addressed by increasing adapt delta to 0.99 and tree depth to 12, remaining divergences were addressed by simplifying the model structure.

The models used for the interaction analyses is as follows:

$$Global_t \sim Local_{t-2} * (Depression_{t-2} + Depression\_mean) + Global_{t-2} + Age + Gender + Education + (1 + Local_{t-2} + Global_{t-2} + Depression_{t-2} + Local_{t-2} : PlayerID : AssessmentDay) \tag{5}$$

$$disc \sim 0 + Depression\_mean + Age + (1|PlayerID) + (1|PlayerID : AssessmentDay)$$

$$Global_t \sim Local_t * (Depression_t + Depression\_mean) + Global_{t-2} + Age + Gender + Education + (1 + Local_t + Global_t + Depression_t + Local_t : Depression_t + Local_t : Depression\_mean||PlayerID) + (1|PlayerID : AssessmentDay) \tag{6}$$

$$disc \sim 0 + Depression\_mean + Age + (1|PlayerID) + (1|PlayerID : AssessmentDay)$$

$$Global_t \sim Local\_mean * Depression\_mean + Age + Gender + Education + (1|PlayerID) + (1|PlayerID : AssessmentDay) \tag{7}$$

$$disc \sim 0 + Depression\_mean + Age + (1|PlayerID) + (1|PlayerID : AssessmentDay)$$

The *disc* parameters were used to model variability in the latent thresholds used for the ratings. We omit the intercept from the *disc* model to ensure that the standard deviation of the latent scale at the intercept is set to one, which aids identifiability (Bürkner & Vuorre, 2019). For the Bayesian analyses, a one-sided test in the hypothesised direction was done. If 0 is located within the 95% credible interval for a one-sided hypothesis test, the parameter values are rejected as there is insufficient evidence in support of the interaction. We reported the 90% equal-tailed credible interval for symmetry. We adopt Stefan et al. (2019) Bayes factor thresholds to classify evidence strength.

Bayesian moderated mediation models

To investigate the mechanism underlying a significant interaction we observed between lagged local confidence and trait depression, we conducted a moderated mediation analysis (Hayes, 2015). Specifically, we tested whether trait depression would significantly impair the within-person autoregression of local confidence (**Path a**; Eq. 8) and subsequently impact the contemporaneous integration of local to global confidence, while accounting for the remaining direct effect of lagged local confidence (**Path b and c'**; Eq. 9). All predictor variables were scaled across the sample, but the predictors and outcome for path a (local confidence) were also within-person centred.

A Bayesian multilevel model including **Path a** and **Path b and c'** was used. For the **Path a** model, unexplained variability in within-person confidence between participants was additionally modelled with parameter *sigma*. Pareto-Smoothed Importance Sampling (PSIS-LOO) cross-validation was used to evaluate model fit for **Path a**, and the model providing the best fit and parsimony was selected. See Supplementary Note 2 for details.

$$\begin{aligned} Local_t \sim & \mathbf{Local_{t-2}}*\mathbf{Depression\_mean} + Depression_{t-2} + Global_{t-2} \\ & + Age + Gender + Education \\ & + (1 + Local_{t-2} + Depression_{t-2} + Global_{t-2}||PlayerID) - \mathbf{Path\ a} \end{aligned} \tag{8}$$

$$sigma \sim 1 + (1|PlayerID)$$

$$\begin{aligned} Global_t \sim & \mathbf{Local_{t-2}}*\mathbf{Depression\_mean} + Depression_{t-2} + Global_{t-2} \\ & + \mathbf{Local_t} + Age + Gender + Education + (1 + Local_{t-2} + Depression_{t-2} \\ & + Global_{t-2} + Local_{t-2} : Depression\_mean||PlayerID) \\ & + (1|PlayerID : AssessmentDay) - \mathbf{Path\ b\ and\ c'} \end{aligned} \tag{9}$$

$$disc \sim 0 + Depression\_mean + age\_s + (1 + PlayerID) + (1 + PlayerID : AssessmentDay)\_\_\_\_\_\_$$

The index of moderated mediation was calculated using the product of the posteriors of **Path a** and **Path b** at ±1 SD of trait depression/symptoms. Similarly, a one-sided hypothesis test in the direction of the hypothesised interaction and moderated mediation findings was done to determine whether there was sufficient evidence supporting our hypothesis. If 0 is located within the 95% credible interval for a one-sided hypothesis test, we determined that there was insufficient evidence in support of the moderated mediation process. We reported the 90% equal-tailed credible interval for symmetry. We adopt Stefan et al., (2019) Bayes factor thresholds to classify evidence strength.

Computational model

To investigate distortions in the mechanisms underlying under-confidence in trait depression and its symptoms, we applied a recently developed computational drift-diffusion model (DDM) of confidence (Katyal & Fleming, 2026). The model builds on prior work modelling choice, decisional and confidence reaction times, and confidence ratings on a two-alternative forced choice task (Desender et al., 2022; Hellmann et al., 2023; Pleskac & Busemeyer, 2010), and was extended to account for distortions in confidence formation during post-decisional evidence accumulation (Katyal & Fleming, 2026). According to traditional DDM models, in each trial of a two-choice task, participants accumulate evidence (*e*) noisily over time at a drift-rate (*v*) until one of two decision boundaries is reached, triggering the corresponding response. After an initial non-decision time (*Ter*), evidence begins accumulating from a starting point (*z*), which is a proportion of the distance between the two decision boundaries (*a*). The reaction time taken for a trial is indexed by the sum of the nondecision time and decision time (*Td*) taken to reach the boundary. We can model the evidence accumulation process at each time step (Δt) as:

$$\Delta e = v^{*} \Delta e + \sigma^{*} N\left(0, \sqrt{\Delta t}\right) \tag{10}$$

The noise (σ) is set to 1 for identifiability purposes (Desender et al., 2022). The At used was .001.

In post-decisional DDM, evidence continues to accumulate after the initial boundary has been reached, accounting for the correspondence between decision accuracy and confidence (Pleskac & Busemeyer, 2010). A ratio of the evidence accumulated after the decision (v_{pd}) over those accumulated during the initial decision process (v) can index a bias-free measure of metacognitive efficiency (v_{ratio}) that accounts for speed-accuracy tradeoff (Desender et al., 2022). This is analogous to M_{ratio} which reflects the ratio of sensitivity of confidence to decision accuracy (*meta-d'*) vs first-order perceptual sensitivity (*d'*) within signal detection frameworks. Similar to the M_{ratio}, a V_{ratio} of 1 implies that the amount of evidence used to estimate confidence (post-decision drift rate) is equal to the evidence used to make the perceptual decision (drift rate), i.e. perfect metacognitive efficiency. However, unlike M_{ratio}, a negative V_{ratio} implies that evidence used to estimate confidence is in the opposite direction of correct evidence i.e., overconfidence in wrong choices and under-confidence in correct choices.

Katyal and Fleming (2026) extended this model by accounting for distortions in post-decisional drift-rate (v_{bias}), which is meant to index a general bias for or against the decision made. We can model the post-decisional drift-rate for each decision in a trial (d_n) as:

$$v_{pd} = v\ ^{*}\ v_{ratio} + d_n{}^{*} V_{bias} \tag{11}$$

Furthermore, unlike M_{ratio} the updated model fit allows V_{ratio} to have positive or negative values. Positive values imply that post-decision evidence accumulates towards higher confidence when the decision is correct and towards lower confidence when the decision is incorrect; negative values imply that post-decision evidence accumulates towards lower confidence when the decision is correct and towards higher confidence when the decision is correct.

Following Katyal and Fleming (2026), we modelled the post-decisional evidence accumulation using a Wiener process that samples empirical confidence response times (rt_{conf}) distributions for each participant:

$$e_{conf} = v_{pd}{}^{*} rt_{conf} + N\left(0, \sqrt{rt_{conf}}\right) \tag{12}$$

The evidence accumulated (e_{conf}), which is an unbounded variable, is then subsequently transformed into a bounded probabilistic estimate of confidence. During this process, parameters indexing the threshold set for confidence estimates (a_{bias}) and the rate by which evidence integrates into confidence (m_{bias}) are introduced:

$$Conf = \frac{1}{1+\exp^{-m_{bias}*(e_{conf}-a_{bias})}} \tag{13}$$

Using differential evolution optimisation in R with the DEOptim package (Mullen et al., 2009), we estimated seven free parameters (v, a, Ter, z, v_{bias}, a_{bias}, m_{bias}, V_{ratio}) by minimising the chi-squared error function as noted in Katyal and Fleming (2026). However, we further binned the responses at each quantile by their stimuli to improve recovery of parameter z:

$$x^2 = \sum_{resp}\sum_{stin}\sum_{cl}\frac{(oRT_{resp.stim,cl}-sRT_{resp,stim,cl})^2}{sRT_{resp,stim,cl}} + \sum_{cl}\sum_{ac}\frac{\left(oConf^{f}{}_{cl,ac}-sConf^{f}{}_{cl,ac}\right)^2}{sConf^{f}{}_{cl,ac}} + \sum_{q}\sum_{ac}\frac{\left(oConf^{s}{}_{cl,ac}-sConf^{s}{}_{cl,ac}\right)^2}{sConf^{s}{}_{cl,ac}} \tag{14}$$

$oRT_{resp,stim,cl}$ and $sRT_{resp,stim,cl}$ denotes the observed and simulated reaction times for each response chosen (resp; i.e., left or right), stimuli (stim; i.e., correct stimuli are on the left or right), and confidence level (cl; 1-6) respectively. $oConf^{f}{}_{cl,ac}$ and $sConf^{f}{}_{cl,ac}$ denotes confidence responses that were faster than their respective rt_{conf}; $oConf^{S}{}_{cl,ac}$ and $sConf^{S}{}_{cl,ac}$ denotes confidence responses that were faster than their respective rt_{conf}. The model was fit separately per participant by aggregating their metacognition assessments across all time points as a single dataset.

The retrieved parameters were z-scored across the whole sample before statistical analyses. We excluded parameter estimates that were larger than 3 standard deviations from the mean, reducing the number of samples to N = 154. To test whether between-person associations between depression and the relevant computational parameters were significant, the following model was used:

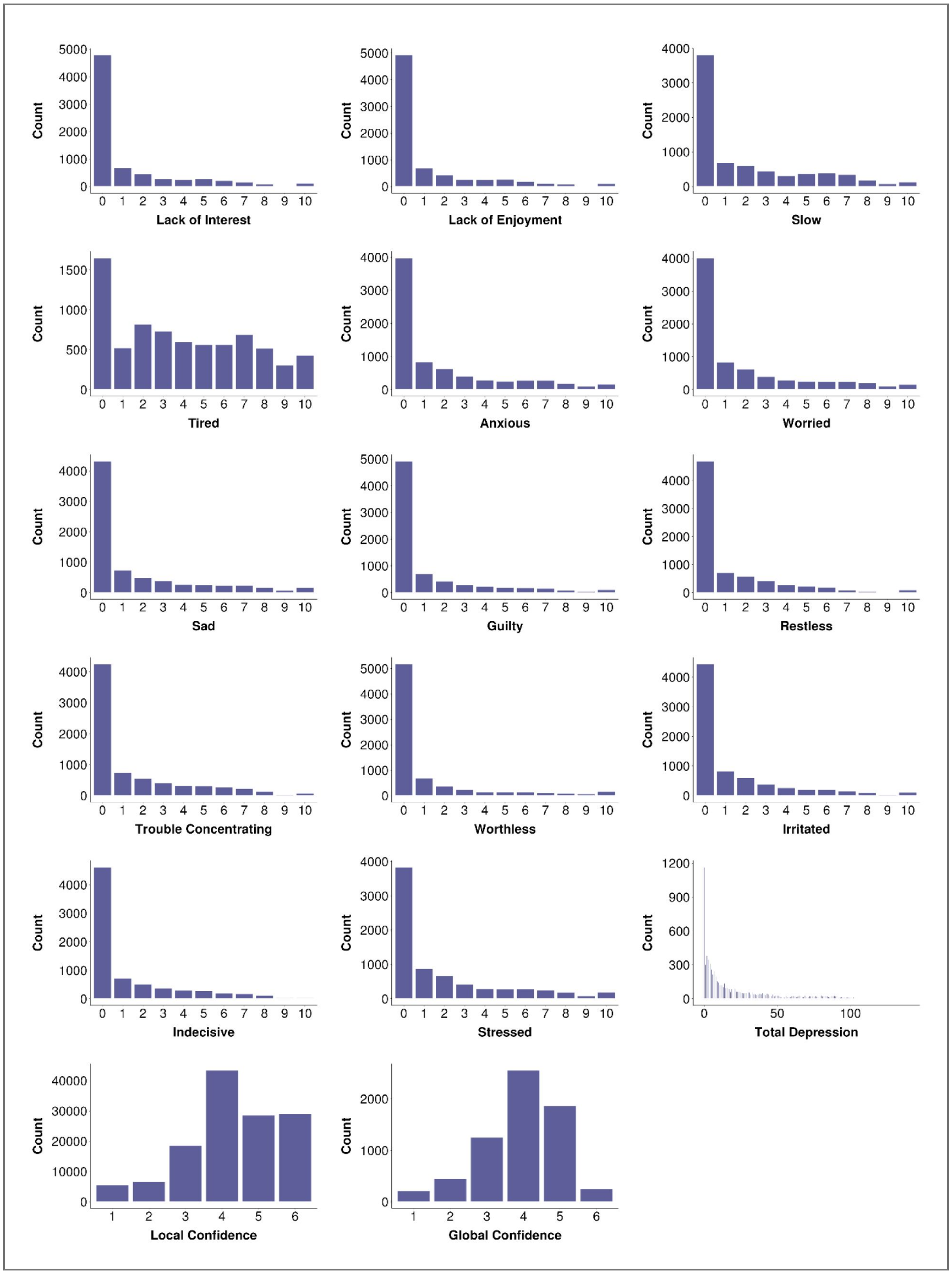
$$Depression \sim v + a + ter + z + v_{bias} + a_{bias} + m_{bias} + V_{ratio} \tag{15}$$

Parameter recovery was conducted using 100 sets of randomly sampled parameter values with 560 trials each and was found to be good to excellent (0.8 < r < 0.99, Ps < .05). See Supplementary Figure 2 for details.

Data availability

All supplementary information, processed data, code, and statistical output is openly accessible at *https://osf.io/5m76h/*.

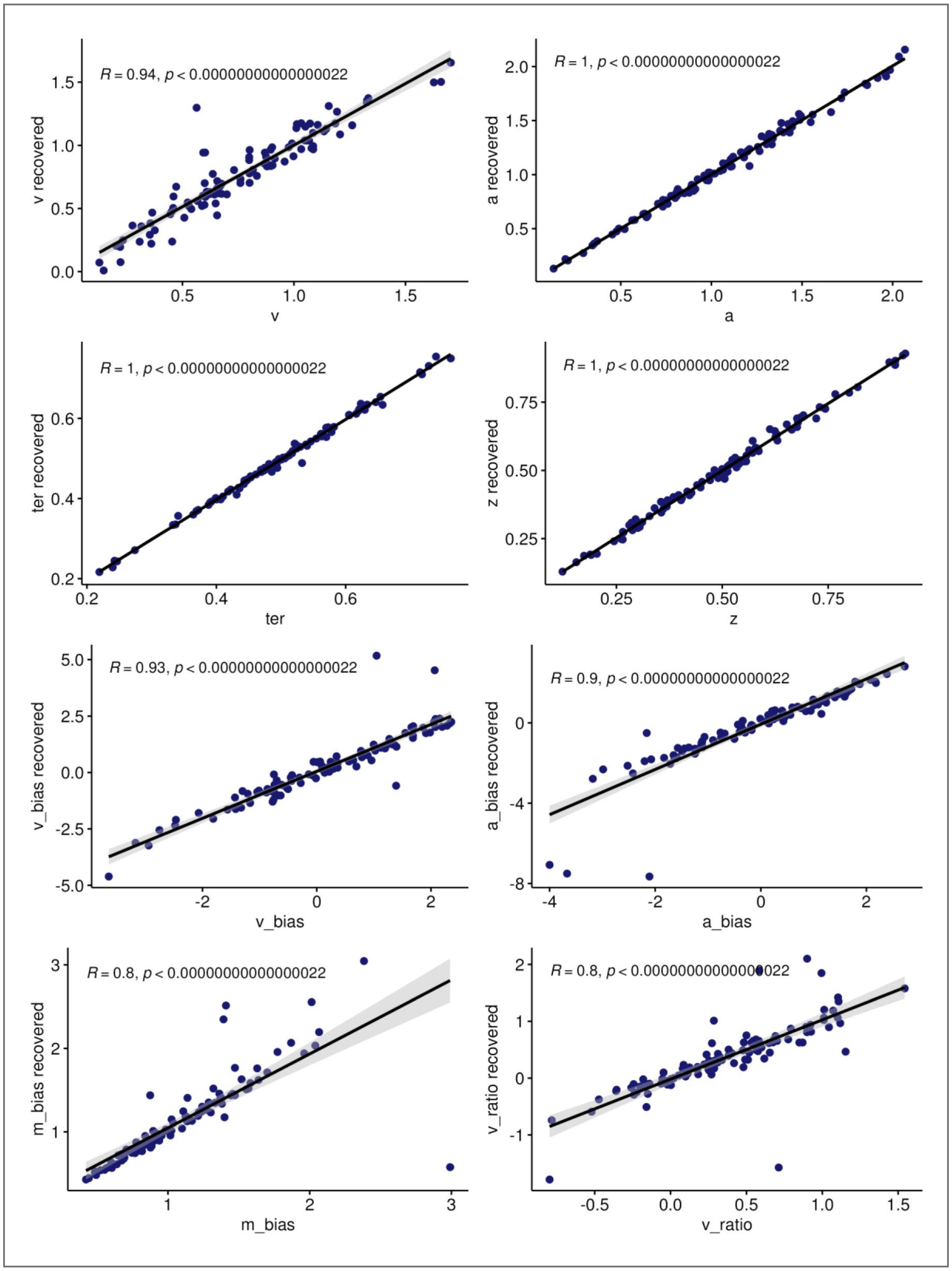
Supplementary Materials



Supplementary Figure 1. Mean EMA Responses and Confidence Ratings

Supplementary Table 1. Between-person EMA Associations with Local Confidence

Items	Partial Correlation	P-BH Resid Forward -> Resid Reverse	P-BH Resid Forward -> Resid Reverse
Lack of Interest	-0.230	0.008	0.008
Lack of Enjoyment	-0.241	0.006	0.007
Slow	-0.206	0.014	0.015



Supplementary Figure 2. **Parameter recovery of simulated data generated by randomly drawn parameters**

Tired	-0.121	0.124	0.154
Anxious	**-0.224**	**0.007**	**0.012**
Worried	**-0.260**	**0.004**	**0.007**
Sad	**-0.260**	**0.004**	**0.007**
Guilty	**-0.170**	**0.036**	**0.029**
Restless	**-0.186**	**0.022**	**0.029**
Trouble Concentrating	**-0.263**	**0.004**	**0.007**
Worthless	**-0.207**	**0.014**	**0.016**
Irritated	**-0.204**	**0.015**	**0.016**
Indecisive	**-0.192**	**0.022**	**0.016**
Stressed	**-0.238**	**0.006**	**0.009**
Total Depression	**-0.239**	**0.002**	**0.002**

Supplementary Table 2. Between-person EMA Associations with Global Confidence

Items	Partial Correlation	P-BH Resid Forward -> Resid Reverse	P-BH Resid Forward -> Resid Reverse
Lack of Interest	**-0.232**	**0.005**	**0.006**
Lack of Enjoyment	**-0.239**	**0.005**	**0.006**
Slow	**-0.226**	**0.005**	**0.006**
Tired	-0.133	0.084	0.118
Anxious	**-0.241**	**0.005**	**0.006**
Worried	**-0.274**	**0.002**	**0.003**
Sad	**-0.285**	**0.002**	**0.003**
Guilty	**-0.203**	**0.011**	**0.010**
Restless	**-0.200**	**0.011**	**0.017**
Trouble Concentrating	**-0.294**	**0.002**	**0.002**
Worthless	**-0.221**	**0.006**	**0.008**
Irritated	**-0.231**	**0.005**	**0.006**
Indecisive	**-0.248**	**0.004**	**0.005**
Stressed	**-0.239**	**0.005**	**0.006**
Total Depression	**-0.262**	**0.001**	**0.001**

Supplementary Table 3. Contemporaneous EMA Associations with Local Confidence

Items	Partial Correlation	P-BH Resid Forward -> Resid Reverse	P-BH Resid Forward -> Resid Reverse
Lack of Interest	0.000	0.980	0.926
Lack of Enjoyment	-0.024	0.742	0.559
Slow	0.005	0.956	0.906
Tired	-0.037	0.508	0.467
Anxious	-0.056	0.508	0.280
Worried	-0.023	0.630	0.559
Sad	-0.031	0.508	0.559
Guilty	-0.005	0.956	0.888
Restless	0.014	0.922	0.716
Trouble Concentrating	-0.073	0.258	0.072
Worthless	-0.017	0.980	0.559
Irritated	-0.019	0.742	0.716

Indecisive	-0.047	0.508	0.388
Stressed	-0.047	0.508	0.335
Total Depression	-0.036	0.364	0.115

Supplementary Table 4. **Contemporaneous EMA Associations with Global Confidence**

Items	Partial Correlation	P-BH Resid Forward -> Resid Reverse	P-BH Resid Forward -> Resid Reverse
Lack of Interest	-0.030	0.314	0.256
Lack of Enjoyment	-0.073	0.055	0.106
Slow	-0.042	0.112	0.106
Tired	-0.057	0.055	0.062
Anxious	-0.072	0.055	0.062
Worried	-0.035	0.163	0.256
Sad	-0.064	0.055	0.062
Guilty	-0.048	0.131	0.106
Restless	-0.024	0.314	0.522
Trouble Concentrating	-0.070	0.057	0.077
Worthless	-0.028	0.343	0.383
Irritated	-0.034	0.163	0.519
Indecisive	-0.067	0.067	0.106
Stressed	-0.061	0.055	0.098
Total Depression	**-0.081**	**0.016**	**0.029**

Supplementary Table 5. **Lagged EMA Predicting Local Confidence (Forward)**

Items	Predictors	Coefficient	CI Low	CI High	P	P-BH
Total Depression	Total Depression Lag	0.002	-0.042	0.045	0.935	0.935
	Total Depression Mean	**-0.247**	**-0.410**	**-0.092**	**0.002**	**0.002**
	Confidence Autoregression	**0.293**	**0.255**	**0.330**	**< .001**	**< .001**
Anxious	Anxious Lag	0.014	-0.014	0.042	0.33	0.786
	Anxious Mean	**-0.258**	**-0.426**	**-0.089**	**0.003**	**0.007**
	Confidence Autoregression	**0.293**	**0.256**	**0.331**	**< .001**	**< .001**
Guilty	Guilty Lag	-0.024	-0.054	0.006	0.115	0.786
	Guilty Mean	**-0.188**	**-0.360**	**-0.017**	**0.033**	**0.036**
	Confidence Autoregression	**0.293**	**0.255**	**0.330**	**< .001**	**< .001**
Indecisive	Indecisive Lag	0.013	-0.019	0.044	0.428	0.786
	Indecisive Mean	**-0.196**	**-0.357**	**-0.035**	**0.018**	**0.022**
	Confidence Autoregression	**0.294**	**0.256**	**0.331**	**< .001**	**< .001**
Irritated	Irritated Lag	0.009	-0.017	0.034	0.503	0.786
	Irritated Mean	**-0.236**	**-0.415**	**-0.057**	**0.011**	**0.015**
	Confidence Autoregression	**0.293**	**0.255**	**0.331**	**< .001**	**< .001**
Lack of Enjoyment	Lack of Enjoyment Lag	0.009	-0.022	0.040	0.558	0.786
	Lack of Enjoyment Mean	**-0.257**	**-0.420**	**-0.094**	**0.002**	**0.006**
	Confidence Autoregression	**0.293**	**0.256**	**0.331**	**< .001**	**< .001**
Lack of Interest	Lack of Interest Lag	-0.018	-0.049	0.012	0.241	0.786
	Lack of Interest Mean	**-0.243**	**-0.408**	**-0.079**	**0.004**	**0.008**
	Confidence Autoregression	**0.293**	**0.256**	**0.330**	**< .001**	**< .001**
Restless	Restless Lag	-0.015	-0.040	0.011	0.253	0.786
	Restless Mean	**-0.217**	**-0.396**	**-0.038**	**0.019**	**0.022**
	Confidence Autoregression	**0.293**	**0.255**	**0.330**	**< .001**	**< .001**
Sad	Sad Lag	0.001	-0.026	0.028	0.948	0.948
	Sad Mean	**-0.283**	**-0.447**	**-0.119**	**0.001**	**0.004**
	Confidence Autoregression	**0.293**	**0.255**	**0.331**	**< .001**	**< .001**
Slow	Slow Lag	-0.014	-0.043	0.014	0.331	0.786
	Slow Mean	**-0.225**	**-0.389**	**-0.061**	**0.008**	**0.014**
	Confidence Autoregression	**0.292**	**0.255**	**0.330**	**< .001**	**< .001**
Stressed	Stressed Lag	-0.013	-0.039	0.014	0.353	0.777
	Stressed Mean	**-0.270**	**-0.442**	**-0.097**	**0.003**	**0.006**
	Confidence Autoregression	**0.292**	**0.255**	**0.330**	**< .001**	**< .001**
Tired	Tired Lag	0.005	-0.019	0.029	0.678	0.791
	Tired Mean	-0.145	-0.329	0.039	0.124	0.124
	Confidence Autoregression	**0.293**	**0.256**	**0.331**	**< .001**	**< .001**
Trouble Concentrating	Trouble Concentrating Lag	-0.004	-0.031	0.024	0.791	0.851
	Trouble Concentrating Mean	**-0.293**	**-0.460**	**-0.125**	**0.001**	**0.004**
	Confidence Autoregression	**0.293**	**0.255**	**0.330**	**< .001**	**< .001**
Worried	Worried Lag	-0.009	-0.037	0.020	0.562	0.786
	Worried Mean	**-0.297**	**-0.464**	**-0.131**	**0.001**	**0.004**
	Confidence Autoregression	**0.293**	**0.255**	**0.330**	**< .001**	**< .001**

Worthless	Worthless Lag	0.008	-0.027	0.042	0.667	0.791
	Worthless Mean	**-0.229**	**-0.398**	**-0.060**	**0.009**	**0.014**
	Confidence Autoregression	**0.293**	**0.255**	**0.331**	**< .001**	**< .001**

Supplementary Table 6. **Lagged EMA Predicting Global Confidence (Forward)**

Items	Predictors	Coefficient	CI Low	CI High	P	P-BH
Total Depression	Depression Lag	-0.015	-0.083	0.054	0.677	0.677
	Depression Mean	**-0.242**	**-0.381**	**-0.103**	**0.001**	**0.001**
	Confidence Autoregression	**0.108**	**0.068**	**0.148**	**< .001**	**< .001**
Anxious	Depression Lag	0.006	-0.039	0.050	0.805	0.86
	Depression Mean	**-0.248**	**-0.399**	**-0.097**	**0.002**	**0.005**
	Confidence Autoregression	**0.109**	**0.068**	**0.149**	**< .001**	**< .001**
Guilty	Depression Lag	-0.043	-0.090	0.005	0.078	0.315
	Depression Mean	**-0.203**	**-0.357**	**-0.049**	**0.011**	**0.011**
	Confidence Autoregression	**0.107**	**0.066**	**0.147**	**< .001**	**< .001**
Indecisive	Depression Lag	0.010	-0.039	0.059	0.694	0.86
	Depression Mean	**-0.228**	**-0.371**	**-0.085**	**0.002**	**0.005**
	Confidence Autoregression	**0.109**	**0.069**	**0.149**	**< .001**	**< .001**
Irritated	Depression Lag	-0.011	-0.051	0.029	0.586	0.86
	Depression Mean	**-0.240**	**-0.400**	**-0.080**	**0.004**	**0.005**
	Confidence Autoregression	**0.108**	**0.068**	**0.148**	**< .001**	**< .001**
Lack of Enjoyment	Depression Lag	-0.002	-0.051	0.046	0.926	0.86
	Depression Mean	**-0.231**	**-0.377**	**-0.084**	**0.002**	**0.005**
	Confidence Autoregression	**0.108**	**0.068**	**0.148**	**< .001**	**< .001**
Lack of Interest	Depression Lag	-0.008	-0.057	0.041	0.751	0.86
	Depression Mean	**-0.222**	**-0.369**	**-0.074**	**0.004**	**0.005**
	Confidence Autoregression	**0.108**	**0.068**	**0.148**	**< .001**	**< .001**
Restless	Depression Lag	-0.027	-0.067	0.013	0.193	0.501
	Depression Mean	**-0.212**	**-0.373**	**-0.052**	**0.01**	**0.011**
	Confidence Autoregression	**0.108**	**0.067**	**0.148**	**< .001**	**< .001**
Sad	Depression Lag	-0.004	-0.046	0.039	0.87	0.86
	Depression Mean	**-0.281**	**-0.428**	**-0.135**	**< .001**	**0.002**
	Confidence Autoregression	**0.108**	**0.068**	**0.148**	**< .001**	**< .001**
Slow	Depression Lag	-0.047	-0.092	-0.002	0.042	0.273
	Depression Mean	**-0.225**	**-0.371**	**-0.078**	**0.003**	**0.005**
	Confidence Autoregression	**0.106**	**0.066**	**0.146**	**< .001**	**< .001**
Stressed	**Depression Lag**	**-0.063**	**-0.105**	**-0.021**	**0.003**	**0.045**
	Depression Mean	**-0.248**	**-0.402**	**-0.095**	**0.002**	**0.005**
	Confidence Autoregression	**0.105**	**0.065**	**0.145**	**< .001**	**< .001**
Tired	Depression Lag	-0.003	-0.040	0.035	0.884	0.86
	Depression Mean	-0.147	-0.312	0.019	0.084	0.084
	Confidence Autoregression	**0.108**	**0.068**	**0.148**	**< .001**	**< .001**
Trouble Concentrating	Depression Lag	-0.037	-0.080	0.007	0.097	0.315
	Depression Mean	**-0.294**	**-0.444**	**-0.145**	**< .001**	**0.002**
	Confidence Autoregression	**0.105**	**0.065**	**0.145**	**< .001**	**< .001**
Worried	Depression Lag	-0.015	-0.060	0.031	0.525	0.6
	Depression Mean	**-0.280**	**-0.429**	**-0.130**	**< .001**	**0.002**
	Confidence Autoregression	**0.108**	**0.068**	**0.148**	**< .001**	**< .001**

Worthless	Depression Lag	-0.030	-0.085	0.023	0.266	0.576
	Depression Mean	**-0.220**	**-0.372**	**-0.068**	**0.005**	**0.006**
	Confidence Autoregression	**0.108**	**0.068**	**0.148**	**< .001**	**< .001**

Supplementary Table 7. **Lagged Local Confidence Predicting EMA (Reverse)**

Items	Predictors	Coefficient	CI Low	CI High	P	P-BH
Total Depression	Confidence Lag	-0.008	-0.042	0.026	0.65	0.65
	Confidence Mean	**-0.230**	**-0.375**	**-0.086**	**0.002**	**0.002**
	Depression Autoregression	**0.046**	**0.006**	**0.086**	**0.023**	**0.023**
Anxious	Confidence Lag	0.002	-0.051	0.055	0.927	0.981
	Confidence Mean	**-0.191**	**-0.326**	**-0.057**	**0.006**	**0.012**
	Depression Autoregression	**0.229**	**0.189**	**0.269**	**< .001**	**< .001**
Guilty	Confidence Lag	-0.010	-0.059	0.038	0.672	0.975
	Confidence Mean	**-0.151**	**-0.284**	**-0.018**	**0.027**	**0.029**
	Depression Autoregression	**0.217**	**0.178**	**0.256**	**< .001**	**< .001**
Indecisive	Confidence Lag	-0.015	-0.064	0.033	0.532	0.945
	Confidence Mean	**-0.189**	**-0.335**	**-0.042**	**0.013**	**0.016**
	Depression Autoregression	**0.128**	**0.088**	**0.168**	**< .001**	**< .001**
Irritated	Confidence Lag	-0.009	-0.067	0.049	0.766	0.975
	Confidence Mean	**-0.173**	**-0.306**	**-0.039**	**0.013**	**0.016**
	Depression Autoregression	**0.149**	**0.111**	**0.188**	**< .001**	**< .001**
Lack of Enjoyment	Confidence Lag	-0.008	-0.056	0.040	0.741	0.975
	Confidence Mean	**-0.225**	**-0.367**	**-0.084**	**0.002**	**0.007**
	Depression Autoregression	**0.133**	**0.094**	**0.172**	**< .001**	**< .001**
Lack of Interest	Confidence Lag	-0.049	-0.098	-0.001	0.045	0.317
	Confidence Mean	**-0.216**	**-0.355**	**-0.077**	**0.003**	**0.008**
	Depression Autoregression	**0.115**	**0.075**	**0.154**	**< .001**	**< .001**
Restless	Confidence Lag	-0.070	-0.127	-0.013	0.015	0.217
	Confidence Mean	**-0.154**	**-0.289**	**-0.020**	**0.026**	**0.029**
	Depression Autoregression	**0.208**	**0.169**	**0.246**	**< .001**	**< .001**
Sad	Confidence Lag	-0.017	-0.073	0.038	0.54	0.945
	Confidence Mean	**-0.237**	**-0.376**	**-0.097**	**0.001**	**0.007**
	Depression Autoregression	**0.227**	**0.187**	**0.266**	**< .001**	**< .001**
Slow	Confidence Lag	-0.044	-0.099	0.008	0.093	0.436
	Confidence Mean	**-0.186**	**-0.324**	**-0.049**	**0.009**	**0.015**
	Depression Autoregression	**0.168**	**0.129**	**0.208**	**< .001**	**< .001**
Stressed	Confidence Lag	-0.025	-0.079	0.030	0.378	0.935
	Confidence Mean	**-0.203**	**-0.338**	**-0.068**	**0.004**	**0.009**
	Depression Autoregression	**0.191**	**0.153**	**0.230**	**< .001**	**< .001**
Tired	Confidence Lag	-0.001	-0.062	0.061	0.981	0.981
	Confidence Mean	-0.098	-0.231	0.036	0.154	0.154
	Depression Autoregression	**0.120**	**0.081**	**0.159**	**< .001**	**< .001**
Trouble Concentrating	Confidence Lag	-0.023	-0.078	0.031	0.402	0.937
	Confidence Mean	**-0.233**	**-0.368**	**-0.099**	**0.001**	**0.007**
	Depression Autoregression	**0.139**	**0.099**	**0.179**	**< .001**	**< .001**
Worried	Confidence Lag	-0.037	-0.089	0.014	0.156	0.547
	Confidence Mean	**-0.223**	**-0.359**	**-0.088**	**0.001**	**0.007**
	Depression Autoregression	**0.194**	**0.155**	**0.233**	**< .001**	**< .001**

Worthless	Confidence Lag	-0.002	-0.045	0.041	0.929	0.981
	Confidence Mean	**-0.184**	**-0.323**	**-0.045**	**0.01**	**0.016**
	Depression Autoregression	**0.103**	**0.064**	**0.143**	**< .001**	**< .001**

Supplementary Table 8. Lagged Global Confidence Predicting EMA (Reverse)

Items	Predictors	Coefficient	CI Low	CI High	P	P-BH
Total Depression	Confidence Lag	-0.011	-0.034	0.012	0.354	0.354
	Confidence Mean	**-0.281**	**-0.443**	**-0.119**	**0.001**	**0.001**
	Depression Autoregression	**0.045**	**0.005**	**0.085**	**0.027**	**0.027**
Anxious	Confidence Lag	0.004	-0.032	0.040	0.841	0.781
	Confidence Mean	**-0.234**	**-0.385**	**-0.084**	**0.003**	**0.006**
	Depression Autoregression	**0.229**	**0.189**	**0.269**	**< .001**	**< .001**
Guilty	Confidence Lag	-0.007	-0.040	0.026	0.669	0.724
	Confidence Mean	**-0.202**	**-0.350**	**-0.053**	**0.009**	**0.01**
	Depression Autoregression	**0.217**	**0.178**	**0.256**	**< .001**	**< .001**
Indecisive	Confidence Lag	-0.005	-0.038	0.028	0.762	0.762
	Confidence Mean	**-0.268**	**-0.430**	**-0.105**	**0.002**	**0.005**
	Depression Autoregression	**0.128**	**0.088**	**0.168**	**< .001**	**< .001**
Irritated	Confidence Lag	-0.020	-0.059	0.019	0.316	0.457
	Confidence Mean	**-0.222**	**-0.371**	**-0.072**	**0.004**	**0.007**
	Depression Autoregression	**0.149**	**0.110**	**0.188**	**< .001**	**< .001**
Lack of Enjoyment	Confidence Lag	-0.009	-0.042	0.024	0.586	0.693
	Confidence Mean	**-0.247**	**-0.406**	**-0.088**	**0.003**	**0.006**
	Depression Autoregression	**0.132**	**0.093**	**0.172**	**< .001**	**< .001**
Lack of Interest	Confidence Lag	-0.029	-0.061	0.004	0.086	0.311
	Confidence Mean	**-0.242**	**-0.398**	**-0.086**	**0.003**	**0.006**
	Depression Autoregression	**0.114**	**0.075**	**0.154**	**< .001**	**< .001**
Restless	Confidence Lag	-0.039	-0.076	0.000	0.048	0.311
	Confidence Mean	**-0.188**	**-0.339**	**-0.036**	**0.016**	**0.017**
	Depression Autoregression	**0.207**	**0.169**	**0.246**	**< .001**	**< .001**
Sad	Confidence Lag	-0.020	-0.057	0.018	0.301	0.457
	Confidence Mean	**-0.289**	**-0.445**	**-0.133**	**< .001**	**0.003**
	Depression Autoregression	**0.226**	**0.186**	**0.265**	**< .001**	**< .001**
Slow	Confidence Lag	-0.047	-0.082	-0.012	0.008	0.106
	Confidence Mean	**-0.228**	**-0.382**	**-0.073**	**0.004**	**0.006**
	Depression Autoregression	**0.166**	**0.127**	**0.205**	**< .001**	**< .001**
Stressed	Confidence Lag	-0.022	-0.059	0.015	0.24	0.455
	Confidence Mean	**-0.230**	**-0.381**	**-0.078**	**0.003**	**0.006**
	Depression Autoregression	**0.191**	**0.152**	**0.229**	**< .001**	**< .001**
Tired	Confidence Lag	-0.022	-0.064	0.019	0.294	0.457
	Confidence Mean	-0.120	-0.270	0.030	0.118	0.118
	Depression Autoregression	**0.119**	**0.080**	**0.158**	**< .001**	**< .001**
Trouble Concentrating	Confidence Lag	-0.023	-0.060	0.014	0.229	0.457
	Confidence Mean	**-0.294**	**-0.444**	**-0.144**	**< .001**	**0.002**
	Depression Autoregression	**0.137**	**0.097**	**0.178**	**< .001**	**< .001**

Worried	Confidence Lag	-0.030	-0.064	0.005	0.096	0.311
	Confidence Mean	**-0.268**	**-0.419**	**-0.117**	**0.001**	**0.003**
	Depression Autoregression	**0.193**	**0.154**	**0.232**	**< .001**	**< .001**
Worthless	Confidence Lag	0.010	-0.019	0.039	0.506	0.658
	Confidence Mean	**-0.221**	**-0.377**	**-0.065**	**0.006**	**0.008**
	Depression Autoregression	**0.104**	**0.064**	**0.143**	**< .001**	**< .001**

Supplementary Table 9. **Same-level Contemporaneous Bayesian Interaction Predicting Global Confidence**

Items	Coefficient	CI Low	CI High	Evidence Ratio	Posterior Probability
Total Depression	0.268	-0.071	0.607	9.485	0.905
Anxious	0.197	-0.031	0.429	11.719	0.921
Guilty	**0.301**	**0.036**	**0.565**	**31.922**	**0.970**
Indecisive	0.217	-0.016	0.450	14.960	0.937
Irritated	0.164	-0.052	0.379	8.662	0.896
Lack of Enjoyment	0.171	-0.087	0.430	6.375	0.864
Lack of Interest	0.144	-0.102	0.396	4.817	0.828
Restless	0.064	-0.166	0.293	2.094	0.677
Sad	0.027	-0.205	0.258	1.365	0.577
Slow	0.005	-0.237	0.241	1.084	0.520
Stressed	0.182	-0.049	0.412	9.492	0.905
Tired	-0.009	-0.203	0.187	0.877	0.467
Trouble Concentrating	**0.463**	**0.231**	**0.699**	**1,141.857**	**0.999**
Worried	0.033	-0.215	0.279	1.389	0.581
Worthless	**0.416**	**0.138**	**0.698**	**161.437**	**0.994**

Supplementary Table 10. **Same-level Trait Bayesian Interaction Predicting Global Confidence**

Items	Coefficient	CI Low	CI High	Evidence Ratio	Posterior Probability
Total Depression	0.024	-0.106	0.155	0.621	0.383
Anxious	0.132	-0.021	0.286	0.086	0.079
Guilty	0.024	-0.119	0.165	0.644	0.392
Indecisive	0.042	-0.109	0.192	0.478	0.323
Irritated	-0.031	-0.181	0.117	1.707	0.631
Lack of Enjoyment	-0.039	-0.177	0.100	2.141	0.682
Lack of Interest	-0.038	-0.177	0.100	2.037	0.671
Restless	-0.043	-0.190	0.103	2.197	0.687
Sad	0.074	-0.066	0.216	0.239	0.193
Slow	0.037	-0.108	0.185	0.513	0.339
Stressed	0.064	-0.084	0.214	0.323	0.244
Tired	0.087	-0.075	0.251	0.228	0.186
Trouble Concentrating	-0.058	-0.207	0.089	2.899	0.744
Worried	0.065	-0.078	0.214	0.305	0.233
Worthless	-0.001	-0.128	0.126	1.034	0.508

Supplementary Table 11. Same-level Lagged Bayesian Interaction Predicting Global Confidence

Items	Coefficient	CI Low	CI High	Evidence Ratio	Posterior Probability
Total Depression	0.185	-0.226	0.599	3.369	0.771
Anxious	-0.001	-0.285	0.288	0.981	0.495
Guilty	**0.305**	**0.003**	**0.609**	**19.672**	**0.952**
Indecisive	**0.463**	**0.136**	**0.795**	**89.909**	**0.989**
Irritated	0.164	-0.087	0.418	6.027	0.858
Lack of Enjoyment	0.219	-0.117	0.553	6.073	0.859
Lack of Interest	0.035	-0.257	0.326	1.380	0.580
Restless	0.121	-0.149	0.394	3.421	0.774
Sad	0.104	-0.174	0.375	2.753	0.734
Slow	0.138	-0.162	0.440	3.536	0.780
Stressed	-0.078	-0.352	0.190	0.466	0.318
Tired	0.038	-0.187	0.282	1.438	0.590
Trouble Concentrating	0.118	-0.146	0.388	3.332	0.769
Worried	0.015	-0.282	0.313	1.154	0.536
Worthless	**0.443**	**0.089**	**0.787**	**49.235**	**0.980**

Supplementary Table 12. Cross-level Contemporaneous Bayesian Interaction Predicting Global Confidence

Items	Coefficient	CI Low	CI High	Evidence Ratio	Posterior Probability
Total Depression	0.078	-0.126	0.288	2.722	0.731
Anxious	0.143	-0.073	0.364	6.197	0.861
Guilty	-0.007	-0.227	0.210	0.928	0.481
Indecisive	0.142	-0.073	0.359	6.146	0.860
Irritated	-0.048	-0.277	0.185	0.574	0.365
Lack of Enjoyment	0.012	-0.212	0.243	1.124	0.529
Lack of Interest	0.041	-0.185	0.276	1.551	0.608
Restless	0.034	-0.198	0.268	1.472	0.595
Sad	0.073	-0.155	0.303	2.316	0.698
Slow	0.101	-0.121	0.328	3.446	0.775
Stressed	0.054	-0.170	0.276	1.922	0.658
Tired	0.204	-0.028	0.438	12.588	0.926
Trouble Concentrating	0.092	-0.134	0.317	3.045	0.753
Worried	0.138	-0.080	0.357	5.647	0.850
Worthless	-0.019	-0.242	0.202	0.797	0.443

Supplementary Table 13. Cross-level Lagged Bayesian Interaction Predicting Global Confidence

Items	Coefficient	CI Low	CI High	Evidence Ratio	Posterior Probability
Total Depression	**-0.279**	**-0.467**	**-0.093**	**156.635**	**0.994**
Anxious	**-0.252**	**-0.463**	**-0.044**	**40.612**	**0.976**
Guilty	**-0.290**	**-0.488**	**-0.094**	**106.744**	**0.991**
Indecisive	**-0.256**	**-0.457**	**-0.058**	**54.459**	**0.982**
Irritated	**-0.313**	**-0.527**	**-0.104**	**152.110**	**0.993**
Lack of Enjoyment	**-0.306**	**-0.501**	**-0.115**	**192.939**	**0.995**

Lack of Interest	**-0.306**	**-0.498**	**-0.113**	**202.822**	**0.995**
Restless	**-0.366**	**-0.576**	**-0.163**	**532.333**	**0.998**
Sad	**-0.222**	**-0.434**	**-0.015**	**24.765**	**0.961**
Slow	**-0.214**	**-0.420**	**-0.005**	**20.813**	**0.954**
Stressed	**-0.313**	**-0.523**	**-0.108**	**148.533**	**0.993**
Tired	**-0.233**	**-0.455**	**-0.011**	**22.375**	**0.957**
Trouble Concentrating	**-0.269**	**-0.468**	**-0.066**	**64.844**	**0.985**
Worried	**-0.257**	**-0.466**	**-0.052**	**50.037**	**0.980**
Worthless	**-0.199**	**-0.395**	**-0.002**	**19.793**	**0.952**

Supplementary Table 14. Trait Depression Moderation on Local Confidence Autoregression

Items	Coefficient	CI Low	CI High	Evidence Ratio	Posterior Probability
Total Depression	**-0.072**	**-0.123**	**-0.022**	**97.462**	**0.990**
Anxious	**-0.071**	**-0.127**	**-0.015**	**49.157**	**0.980**
Guilty	**-0.083**	**-0.137**	**-0.029**	**159.000**	**0.994**
Indecisive	**-0.092**	**-0.144**	**-0.041**	**550.724**	**0.998**
Irritated	**-0.066**	**-0.125**	**-0.007**	**28.520**	**0.966**
Lack of Enjoyment	**-0.080**	**-0.134**	**-0.026**	**131.780**	**0.992**
Lack of Interest	**-0.074**	**-0.129**	**-0.020**	**76.859**	**0.987**
Restless	**-0.067**	**-0.122**	**-0.012**	**42.243**	**0.977**
Sad	**-0.063**	**-0.118**	**-0.007**	**30.311**	**0.968**
Slow	-0.043	-0.099	0.012	8.926	0.899
Stressed	**-0.064**	**-0.120**	**-0.009**	**33.043**	**0.971**
Tired	-0.041	-0.100	0.019	6.713	0.870
Trouble Concentrating	**-0.069**	**-0.127**	**-0.012**	**40.721**	**0.976**
Worried	**-0.061**	**-0.116**	**-0.006**	**28.385**	**0.966**
Worthless	**-0.057**	**-0.112**	**-0.003**	**23.922**	**0.960**

Supplementary Table 15. Index of Moderated Mediation on Global Confidence for Depression at ±1 SD

Items	Coefficient	CI Low	CI High	Evidence Ratio	Posterior Probability
Total Depression	**-0.310**	**-0.533**	**-0.090**	**95.096**	**0.990**
Anxious	**-0.303**	**-0.547**	**-0.063**	**52.601**	**0.981**
Guilty	**-0.354**	**-0.588**	**-0.122**	**161.437**	**0.994**
Indecisive	**-0.393**	**-0.622**	**-0.170**	**437.356**	**0.998**
Irritated	**-0.282**	**-0.537**	**-0.030**	**29.769**	**0.968**
Lack of Enjoyment	**-0.343**	**-0.577**	**-0.112**	**130.148**	**0.992**
Lack of Interest	**-0.318**	**-0.556**	**-0.082**	**73.766**	**0.987**
Restless	**-0.288**	**-0.526**	**-0.050**	**44.911**	**0.978**
Sad	**-0.270**	**-0.510**	**-0.034**	**31.000**	**0.969**
Slow	-0.188	-0.425	0.047	9.359	0.903
Stressed	**-0.272**	**-0.514**	**-0.034**	**30.936**	**0.969**
Tired	-0.173	-0.431	0.083	6.771	0.871
Trouble Concentrating	**-0.300**	**-0.542**	**-0.058**	**47.855**	**0.980**
Worried	**-0.257**	**-0.498**	**-0.021**	**26.142**	**0.963**

Worthless	**-0.242**	**-0.483**	**-0.006**	**20.666**	**0.954**

Supplementary Table 16. Direct Effect of Interaction on Global Confidence (Path c')

Items	Coefficient	CI Low	CI High	Evidence Ratio	Posterior Probability
Total Depression	-0.124	-0.253	0.002	17.790	0.947
Anxious	-0.108	-0.251	0.029	9.420	0.904
Guilty	-0.104	-0.239	0.027	9.516	0.905
Indecisive	-0.086	-0.222	0.048	5.895	0.855
Irritated	-0.117	-0.266	0.029	9.811	0.907
Lack of Enjoyment	-0.105	-0.244	0.030	9.194	0.902
Lack of Interest	-0.106	-0.247	0.031	8.762	0.898
Restless	**-0.201**	**-0.346**	**-0.059**	**99.946**	**0.990**
Sad	-0.090	-0.235	0.051	5.883	0.855
Slow	-0.134	-0.276	0.007	15.780	0.940
Stressed	**-0.150**	**-0.294**	**-0.011**	**25.230**	**0.962**
Tired	-0.122	-0.276	0.032	9.663	0.906
Trouble Concentrating	-0.128	-0.266	0.008	15.377	0.939
Worried	-0.135	-0.276	0.001	18.370	0.948
Worthless	-0.099	-0.231	0.033	8.417	0.894

Supplementary Table 17. Depression Associations with Computational Parameters

Items	Predictors	Coefficient	CI Low	CI High	P	P-BH
Total Depression	v	-0.050	-0.244	0.143	0.612	0.612
	a	-0.239	-0.477	-0.002	0.05	0.05
	ter	-0.101	-0.262	0.061	0.225	0.225
	z	0.105	-0.029	0.240	0.128	0.128
	v-bias	-0.163	-0.374	0.048	0.131	0.131
	a-bias	**0.381**	**0.145**	**0.617**	**0.002**	**0.002**
	m-bias	-0.021	-0.192	0.151	0.815	0.815
	V-ratio	-0.041	-0.198	0.115	0.605	0.605
	Age	**-0.174**	**-0.329**	**-0.019**	**0.029**	**0.029**
	Male	0.265	-0.058	0.588	0.11	0.11
	Non-cisgender	**2.174**	**0.987**	**3.362**	**< .001**	**< .001**
	Prefer	-0.486	-2.106	1.133	0.557	0.557
	Education (University)	-0.247	-0.538	0.045	0.099	0.099

Anxious	v	-0.040	-0.223	0.142	0.666	0.945
	a	-0.215	-0.439	0.008	0.061	0.122
	ter	-0.131	-0.283	0.021	0.093	0.435
	z	0.074	-0.053	0.201	0.253	0.314
	v-bias	-0.132	-0.331	0.066	0.193	0.3
	a-bias	**0.313**	**0.090**	**0.535**	**0.007**	**0.01**
	m-bias	0.003	-0.158	0.165	0.967	0.967
	V-ratio	-0.086	-0.234	0.061	0.254	0.97
	Age	-0.143	-0.289	0.003	0.057	0.078
	Male	0.196	-0.109	0.500	0.209	0.114
	Non-cisgender	**2.132**	**1.013**	**3.251**	**< .001**	**0.001**
	Prefer	-0.367	-1.893	1.159	0.638	0.704
	Education (University)	-0.163	-0.438	0.112	0.247	0.314
Guilty	v	0.006	-0.171	0.183	0.943	0.945
	a	-0.230	-0.447	-0.013	0.04	0.122
	ter	-0.160	-0.308	-0.013	0.035	0.435
	z	0.070	-0.054	0.193	0.269	0.314
	v-bias	-0.073	-0.265	0.119	0.458	0.494
	a-bias	0.214	-0.002	0.430	0.053	0.058
	m-bias	0.017	-0.140	0.174	0.829	0.967
	V-ratio	-0.003	-0.147	0.140	0.965	0.97
	Age	-0.144	-0.285	-0.002	0.048	0.078
	Male	0.166	-0.129	0.462	0.272	0.136
	Non-cisgender	**2.797**	**1.712**	**3.882**	**< .001**	**< .001**
	Prefer	-0.902	-2.382	0.578	0.234	0.704
	Education (University)	-0.246	-0.512	0.021	0.073	0.235
Indecisive	v	-0.026	-0.219	0.168	0.795	0.945
	a	-0.239	-0.477	-0.002	0.05	0.122
	ter	-0.066	-0.228	0.095	0.424	0.59
	z	0.081	-0.054	0.216	0.243	0.314
	v-bias	-0.088	-0.299	0.122	0.413	0.482
	a-bias	**0.316**	**0.080**	**0.551**	**0.01**	**0.013**
	m-bias	-0.006	-0.177	0.166	0.949	0.967
	V-ratio	0.003	-0.154	0.160	0.97	0.97
	Age	-0.152	-0.307	0.002	0.055	0.078
	Male	0.309	-0.014	0.632	0.062	0.08
	Non-cisgender	**1.780**	**0.593**	**2.967**	**0.004**	**0.006**
	Prefer	0.876	-0.743	2.495	0.291	0.704
	Education (University)	-0.206	-0.497	0.085	0.168	0.235

Irritated	v	-0.039	-0.217	0.138	0.664	0.945
	a	-0.181	-0.399	0.037	0.105	0.134
	ter	-0.060	-0.208	0.089	0.431	0.59
	z	0.116	-0.008	0.239	0.069	0.219
	v-bias	-0.211	-0.404	-0.018	0.034	0.285
	a-bias	**0.325**	**0.108**	**0.541**	**0.004**	**0.007**
	m-bias	-0.059	-0.217	0.098	0.462	0.967
	V-ratio	-0.050	-0.194	0.094	0.496	0.97
	Age	-0.137	-0.279	0.005	0.061	0.078
	Male	0.270	-0.027	0.566	0.077	0.08
	Non-cisgender	1.043	-0.047	2.133	0.063	0.063
	Prefer	-0.269	-1.756	1.218	0.723	0.723
	Education (University)	-0.226	-0.493	0.042	0.1	0.235
Lack of Enjoyment	v	-0.102	-0.283	0.079	0.27	0.945
	a	-0.250	-0.472	-0.028	0.028	0.122
	ter	-0.032	-0.183	0.119	0.676	0.676
	z	0.119	-0.007	0.245	0.066	0.219
	v-bias	-0.165	-0.361	0.032	0.103	0.285
	a-bias	**0.354**	**0.133**	**0.574**	**0.002**	**0.005**
	m-bias	-0.086	-0.246	0.075	0.297	0.967
	V-ratio	-0.017	-0.163	0.130	0.823	0.97
	Age	-0.163	-0.308	-0.019	0.028	0.078
	Male	0.308	0.006	0.610	0.047	0.08
	Non-cisgender	**2.082**	**0.973**	**3.192**	**< .001**	**0.001**
	Prefer	-0.357	-1.870	1.156	0.645	0.704
	Education (University)	-0.217	-0.489	0.055	0.12	0.235
Lack of Interest	v	-0.103	-0.283	0.077	0.263	0.945
	a	-0.264	-0.484	-0.043	0.021	0.122
	ter	-0.048	-0.198	0.102	0.531	0.59
	z	0.128	0.003	0.254	0.047	0.219
	v-bias	-0.169	-0.364	0.027	0.094	0.285
	a-bias	**0.358**	**0.139**	**0.578**	**0.002**	**0.005**
	m-bias	-0.082	-0.242	0.077	0.315	0.967
	V-ratio	-0.034	-0.179	0.112	0.651	0.97
	Age	-0.168	-0.312	-0.024	0.024	0.078
	Male	0.287	-0.014	0.587	0.063	0.08
	Non-cisgender	**2.024**	**0.920**	**3.128**	**< .001**	**0.001**
	Prefer	-0.430	-1.936	1.075	0.576	0.704
	Education (University)	-0.205	-0.476	0.066	0.14	0.235

Restless	v	0.014	-0.168	0.196	0.881	0.945
	a	-0.101	-0.325	0.122	0.375	0.375
	ter	-0.047	-0.199	0.105	0.548	0.59
	z	0.109	-0.018	0.236	0.094	0.219
	v-bias	-0.126	-0.324	0.072	0.214	0.3
	a-bias	**0.293**	**0.071**	**0.515**	**0.01**	**0.013**
	m-bias	0.006	-0.155	0.168	0.938	0.967
	V-ratio	-0.004	-0.151	0.144	0.961	0.97
	Age	-0.103	-0.249	0.042	0.167	0.179
	Male	0.212	-0.092	0.516	0.173	0.104
	Non-cisgender	**1.575**	**0.458**	**2.692**	**0.006**	**0.009**
	Prefer	-0.350	-1.873	1.174	0.654	0.704
	Education (University)	-0.375	-0.649	-0.101	0.008	0.113
Sad	v	-0.037	-0.220	0.147	0.697	0.945
	a	-0.196	-0.421	0.029	0.089	0.134
	ter	-0.075	-0.228	0.078	0.337	0.59
	z	0.115	-0.013	0.243	0.079	0.219
	v-bias	-0.151	-0.350	0.049	0.141	0.285
	a-bias	**0.348**	**0.125**	**0.571**	**0.003**	**0.005**
	m-bias	0.017	-0.146	0.179	0.841	0.967
	V-ratio	-0.107	-0.256	0.042	0.16	0.97
	Age	-0.066	-0.212	0.081	0.381	0.381
	Male	0.249	-0.057	0.555	0.113	0.085
	Non-cisgender	1.093	-0.031	2.218	0.059	0.063
	Prefer	-0.560	-2.094	0.974	0.476	0.704
	Education (University)	-0.149	-0.426	0.127	0.29	0.339
Slow	v	-0.051	-0.235	0.134	0.593	0.945
	a	-0.225	-0.451	0.002	0.054	0.122
	ter	-0.055	-0.209	0.099	0.488	0.59
	z	0.103	-0.026	0.231	0.119	0.239
	v-bias	-0.120	-0.321	0.081	0.242	0.308
	a-bias	**0.352**	**0.127**	**0.577**	**0.003**	**0.005**
	m-bias	-0.013	-0.177	0.151	0.876	0.967
	V-ratio	0.012	-0.138	0.161	0.877	0.97
	Age	-0.160	-0.308	-0.013	0.035	0.078
	Male	0.257	-0.051	0.565	0.104	0.085
	Non-cisgender	1.159	0.026	2.292	0.047	0.059
	Prefer	-0.566	-2.111	0.979	0.474	0.704
	Education (University)	-0.199	-0.477	0.079	0.163	0.235

Stressed	v	-0.006	-0.184	0.172	0.945	0.945
	a	-0.162	-0.381	0.056	0.147	0.171
	ter	-0.131	-0.279	0.018	0.086	0.435
	z	0.043	-0.080	0.167	0.493	0.531
	v-bias	-0.156	-0.350	0.037	0.116	0.285
	a-bias	**0.355**	**0.138**	**0.572**	**0.002**	**0.005**
	m-bias	-0.019	-0.177	0.138	0.811	0.967
	V-ratio	-0.051	-0.195	0.093	0.489	0.97
	Age	-0.104	-0.246	0.039	0.155	0.179
	Male	0.229	-0.068	0.526	0.132	0.088
	Non-cisgender	**2.081**	**0.990**	**3.173**	**< .001**	**0.001**
	Prefer	-0.396	-1.885	1.092	0.602	0.704
	Education (University)	-0.127	-0.394	0.141	0.356	0.383
Tired	v	-0.038	-0.214	0.138	0.676	0.945
	a	-0.102	-0.318	0.114	0.355	0.375
	ter	-0.067	-0.213	0.080	0.376	0.59
	z	0.005	-0.118	0.127	0.938	0.938
	v-bias	-0.057	-0.248	0.135	0.562	0.562
	a-bias	0.172	-0.042	0.387	0.118	0.118
	m-bias	0.033	-0.123	0.189	0.679	0.967
	V-ratio	-0.028	-0.170	0.115	0.705	0.97
	Age	-0.170	-0.311	-0.030	0.019	0.078
	Male	-0.036	-0.330	0.258	0.81	0.347
	Non-cisgender	1.072	-0.007	2.152	0.053	0.062
	Prefer	-0.670	-2.143	0.803	0.374	0.704
	Education (University)	-0.193	-0.458	0.072	0.156	0.235
Trouble Concentrating	v	-0.034	-0.213	0.145	0.711	0.945
	a	-0.188	-0.408	0.032	0.096	0.134
	ter	-0.069	-0.218	0.081	0.369	0.59
	z	0.083	-0.042	0.207	0.197	0.314
	v-bias	-0.145	-0.340	0.050	0.147	0.285
	a-bias	**0.378**	**0.160**	**0.596**	**0.001**	**0.005**
	m-bias	-0.039	-0.198	0.120	0.629	0.967
	V-ratio	0.023	-0.122	0.168	0.756	0.97
	Age	-0.195	-0.338	-0.052	0.009	0.078
	Male	0.269	-0.030	0.569	0.08	0.08
	Non-cisgender	**1.667**	**0.568**	**2.767**	**0.003**	**0.006**
	Prefer	-0.367	-1.866	1.133	0.633	0.704
	Education (University)	-0.270	-0.540	0.000	0.052	0.235

Worried	v	-0.044	-0.224	0.137	0.637	0.945
	a	-0.192	-0.413	0.029	0.091	0.134
	ter	-0.102	-0.252	0.048	0.186	0.521
	z	0.082	-0.044	0.207	0.204	0.314
	v-bias	-0.153	-0.349	0.043	0.128	0.285
	a-bias	**0.362**	**0.142**	**0.582**	**0.002**	**0.005**
	m-bias	-0.028	-0.188	0.131	0.729	0.967
	V-ratio	-0.040	-0.186	0.106	0.592	0.97
	Age	-0.142	-0.286	0.002	0.056	0.078
	Male	0.152	-0.149	0.452	0.325	0.15
	Non-cisgender	**2.103**	**0.997**	**3.209**	**< .001**	**0.001**
	Prefer	-0.481	-1.989	1.027	0.533	0.704
	Education (University)	-0.219	-0.491	0.052	0.115	0.235
Worthless	v	-0.075	-0.252	0.103	0.41	0.945
	a	-0.242	-0.459	-0.024	0.031	0.122
	ter	-0.113	-0.261	0.035	0.135	0.472
	z	0.126	0.002	0.249	0.047	0.219
	v-bias	-0.138	-0.331	0.055	0.163	0.285
	a-bias	**0.276**	**0.060**	**0.492**	**0.013**	**0.015**
	m-bias	0.007	-0.150	0.165	0.926	0.967
	V-ratio	-0.079	-0.222	0.065	0.283	0.97
	Age	-0.151	-0.292	-0.009	0.039	0.078
	Male	0.310	0.014	0.606	0.041	0.08
	Non-cisgender	**2.846**	**1.759**	**3.933**	**< .001**	**< .001**
	Prefer	-0.620	-2.102	0.863	0.414	0.704
	Education (University)	-0.084	-0.351	0.182	0.536	0.536

Supplementary Table 18. Demographic and descriptive statistics for the main samples

Variables	Stats	Count	Proportion
Age	54.38 (11.34) [19 - 82]		
Gender			
Female		122	75.31
Male		37	22.84
Non-cisgender		2	1.23
Prefer not to answer		1	0.62
Education			
Sub-university		49	30.25
University		113	69.75
Country			
United Kingdom		129	79.63
Ireland		13	8.02
United States		13	8.02
Other		7	4.32
Language			
English		149	91.98

Other	13	8.02

Supplementary Table 19. Mean Responses in EMA Questionnaires and Confidence

Items	Mean	Median	SD
Lack of Interest	1.31 (2.1)	1.17 (2.17)	0.74 (0.79)
Lack of Enjoyment	1.21 (2.03)	1.06 (2.05)	0.67 (0.78)
Slow	2.01 (2.4)	1.87 (2.63)	1.1 (0.82)
Tired	3.91 (2.65)	3.83 (2.97)	1.66 (0.78)
Anxious	1.82 (2.36)	1.66 (2.56)	1.01 (0.88)
Worried	1.79 (2.36)	1.65 (2.52)	1 (0.85)
Sad	1.62 (2.27)	1.48 (2.51)	1.02 (0.9)
Guilty	1.19 (2.01)	1.06 (2.13)	0.68 (0.77)
Restless	1.23 (1.79)	1.03 (1.9)	0.85 (0.83)
Trouble Concentrating	1.58 (2.11)	1.4 (2.25)	0.92 (0.84)
Worthless	1.12 (2.15)	1.02 (2.21)	0.56 (0.8)
Irritated	1.36 (1.9)	1.12 (2.04)	1 (0.85)
Indecisive	1.33 (2)	1.19 (2.06)	0.72 (0.72)
Stressed	1.88 (2.33)	1.69 (2.51)	1.1 (0.87)
Total Depression	23.43 (27.13)	22.86 (27.57)	7.12 (5.74)
Local Confidence	4.29 (0.89)	4.29 (0.91)	0.32 (0.17)
Global Confidence	3.93 (0.86)	3.95 (0.87)	0.51 (0.25)

Supplementary Note 1. Weakly Informative Priors

Analysis	Parameter	Prior
Interaction Analysis	Intercept	student_t(3, -1.65, 1), coef = 1 student_t(3, -1, 1), coef = 2, student_t(3, 0, 1), coef = 3, student_t(3, 1, 1), coef = 4, student_t(3, 1.65, 1), coef = 5
	b	normal(0, 0.5)
	sd	student_t(3, 0, 0.5)
	b_disc	normal(0, 0.35)
	sd_disc	normal(0, 0.35)
Moderated Mediation Analysis	Intercept (Path a)	normal(0, 0.05)
	b (Path a)	normal(0, 0.5)
	sd (Path a)	student_t(3, 0, 0.5)
	Intercept (Path b)	student_t(3, -1.65, 1), coef = 1 student_t(3, -1, 1), coef = 2, student_t(3, 0, 1), coef = 3, student_t(3, 1, 1), coef = 4, student_t(3, 1.65, 1), coef = 5
	b (Path b)	normal(0, 0.5)
	sd (Path b)	student_t(3, 0, 0.5)
	b_disc (Path b)	normal(0, 0.35)
	sd_disc (Path b)	normal(0, 0.35)

Supplementary Note 2. **PSIS-LOO Cross Validation of Mediation Model of Path a**

Random Term and Sigma	ELPD LOO Estimate (SE)
(1 + MeanConf_s_two_lag_c + GlobalConf_s_two_lag_c + TotalDepression_dt_s_two_lag_c + MeanConf_s_two_lag_c:TotalDepression_dt_s_mean ‖ PlayerID) sigma ~ TotalDepression_dt_s_mean + GlobalConf_s_mean + MeanConf_s_mean + (1 | PlayerID)	-86.1 (72.1)
(1 + MeanConf_s_two_lag_c + GlobalConf_s_two_lag_c + TotalDepression_dt_s_two_lag_c + MeanConf_s_two_lag_c:TotalDepression_dt_s_mean ‖ PlayerID) sigma ~ 1 + (1 | PlayerID)	-86.4 (72.1)
(1 + MeanConf_s_two_lag_c + GlobalConf_s_two_lag_c + TotalDepression_dt_s_two_lag_c ‖ PlayerID) sigma ~ TotalDepression_dt_s_mean + GlobalConf_s_mean + MeanConf_s_mean + (1 | PlayerID)	-86.7 (72.1)
(1 + MeanConf_s_two_lag_c + GlobalConf_s_two_lag_c + TotalDepression_dt_s_two_lag_c ‖ PlayerID) sigma ~ 1 + (1 | PlayerID)	-86.8 (72.0)
(1 + MeanConf_s_two_lag_c + GlobalConf_s_two_lag_c + TotalDepression_dt_s_two_lag_c ‖ PlayerID) sigma ~ TotalDepression_dt_s_mean + (1 | PlayerID)	-87.8 (72.1)
(1 + MeanConf_s_two_lag_c + GlobalConf_s_two_lag_c ‖ PlayerID) sigma ~ 1 + (1 | PlayerID)	-143.1 (72.0)
(1 + MeanConf_s_two_lag_c + GlobalConf_s_two_lag_c ‖ PlayerID) sigma ~ TotalDepression_dt_s_mean + GlobalConf_s_mean + MeanConf_s_mean + (1 | PlayerID)	-143.7 (72.0)
(1 + MeanConf_s_two_lag_c + GlobalConf_s_two_lag_c + MeanConf_s_two_lag_c:TotalDepression_dt_s_mean ‖ PlayerID) sigma ~ 1 + (1 | PlayerID)	-145 (72.0)
(1 + MeanConf_s_two_lag_c + GlobalConf_s_two_lag_c + MeanConf_s_two_lag_c:TotalDepression_dt_s_mean ‖ PlayerID) sigma ~ TotalDepression_dt_s_mean + GlobalConf_s_mean + MeanConf_s_mean + (1 | PlayerID)	-146.7 (72.0)

Additional information

Funding

Funder	Grant reference number	Author
Science Foundation Ireland (SFI)	SFI-19/DP/7213	Claire M Gillan
Science Foundation Ireland (SFI)	21/RC/10294	Paveen Phon Amnuaisuk Vanessa Teckentrup Claire M Gillan

Author ORCID iDs

Sol Fittipaldi: https://orcid.org/0000-0001-7704-3349
Stephen M Fleming: https://orcid.org/0000-0003-0233-4891
Claire M Gillan: https://orcid.org/0000-0001-9065-403X

References

Agnoli S, Zuberer A, Nanni-Zepeda M, McGlinchey RE, Milberg WP, Esterman M, DeGutis J (2023) Depressive Symptoms are Associated with More Negative Global Metacognitive Biases in Combat Veterans, and Biases Covary with Symptom Changes over Time. *Depression and Anxiety* **2023**:2925551 https://doi.org/10.1155/2023/2925551 | PubMed

Bates D, Mächler M, Bolker B, Walker S (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* **67**:1-48 https://doi.org/10.18637/jss.v067.i01

Beck AT (1979) *Cognitive Therapy and the Emotional Disorders* Penguin.

Benjamini Y, Krieger AM, Yekutieli D (2006) Adaptive Linear Step-up Procedures That Control the False Discovery Rate. *Biometrika* **93**:491-507 https://doi.org/10.1093/biomet/93.3.491

Bénon J, Lee D, Hopper W, Verdeil M, Pessiglione M, Vinckier F, Bouret S, Rouault M, Lebouc R, Pezzulo G, *et al.* (2024) The online metacognitive control of decisions. *Communications Psychology* **2**:1-17 https://doi.org/10.1038/s44271-024-00071-y | PubMed

Benwell CSY, Mohr G, Wallberg J, Kouadio A, Ince RAA (2022) Psychiatrically relevant signatures of domain-general decision-making and metacognition in the general population. *Npj Mental Health Research* **1**:1-17 https://doi.org/10.1038/s44184-022-00009-4 | PubMed

Boettiger C, Eddelbuettel D (2017) An Introduction to Rocker: Docker Containers for R. *The R Journal* **9**:527-536 https://doi.org/10.32614/RJ-2017-065

Bringmann LF, Vissers N, Wichers M, Geschwind N, Kuppens P, Peeters F, Borsboom D, Tuerlinckx F (2013) A Network Approach to Psychopathology: New Insights into Clinical Longitudinal Data. *PLOS One* **8**:e60188 https://doi.org/10.1371/journal.pone.0060188 | PubMed

Bürkner P.-C (2017) brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software* **80**:1-28 https://doi.org/10.18637/jss.v080.i01

Bürkner P.-C, Vuorre M (2019) Ordinal Regression Models in Psychology: A Tutorial. *Advances in Methods and Practices in Psychological Science* **2**:77-101 https://doi.org/10.1177/2515245918823199

Culot C, Fantini-Hauwel C, Gevers W (2021) The influence of sad mood induction on task performance and metacognition. *Quarterly Journal of Experimental Psychology* **74**:1605-1614 https://doi.org/10.1177/17470218211004205 | PubMed

Culot C, Lauwers T, Fantini-Hauwel C, Madani Y, Schrijvers D, Morrens M, Gevers W (2022) Contributions of age and clinical depression to metacognitive performance. *medRxiv* 2022.01.27.22269959 https://doi.org/10.1101/2022.01.27.22269959

da Fonseca M, Maffei G, Moreno-Bote R, Hyafil A (2023) Mood and implicit confidence independently fluctuate at different time scales. *Cognitive, Affective, & Behavioral Neuroscience* **23**:142-161 https://doi.org/10.3758/s13415-022-01038-4 | PubMed

Desender K, Vermeylen L, Verguts T (2022) Dynamic influences on static measures of metacognition. *Nature Communications* **13**:4208 https://doi.org/10.1038/s41467-022-31727-0 | PubMed

Epskamp S, Waldorp LJ, Mõttus R, Borsboom D (2018) The Gaussian Graphical Model in Cross-sectional and Time-series Data. *arXiv* https://doi.org/10.48550/arXiv.1609.04156

Fleming SM, Dolan RJ (2012) The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**:1338-1349 https://doi.org/10.1098/rstb.2011.0417 | PubMed

Fox CA, Lee CT, Hanlon AK, Seow TX, Lynch K, Harty S, Richards D, Palacios J, O'Keane V, Stephan KE, *et al.* (2023) An observational treatment study of metacognition in anxious-depression. *eLife* **12**:RP87193 https://doi.org/10.7554/eLife.87193 | PubMed

Fox CA, McDonogh A, Donegan KR, Teckentrup V, Crossen RJ, Hanlon AK, Gallagher E, Rouault M, Gillan CM (2024) Reliable, rapid, and remote measurement of metacognitive bias. *Scientific Reports* **14**:14941 https://doi.org/10.1038/s41598-024-64900-0 | PubMed

Fried EI, Epskamp S, Nesse RM, Tuerlinckx F, Borsboom D (2016) What are "good" depression symptoms? Comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis. *Journal of Affective Disorders* **189**:314-320 https://doi.org/10.1016/j.jad.2015.09.005 | PubMed

Fried EI, Nesse RM (2015) Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR*D study. *Journal of Affective Disorders* **172**:96-102 https://doi.org/10.1016/j.jad.2014.10.010 | PubMed

Gelman A, Vehtari A, Simpson D, Margossian CC, Carpenter B, Yao Y, Kennedy L, Gabry J, Bürkner P.-C, Modrák M (2020) Bayesian Workflow. *arXiv* https://doi.org/10.48550/arXiv.2011.01808

Guggenmos M (2021) Measuring metacognitive performance: Type 1 performance dependence and test-retest reliability. *Neuroscience of Consciousness* **2021**:niab040 https://doi.org/10.1093/nc/niab040 | PubMed

Hayes AF (2015) An Index and Test of Linear Moderated Mediation. *Multivariate Behavioral Research* **50**:1-22 https://doi.org/10.1080/00273171.2014.962683 | PubMed

Hobbs C, Vozarova P, Sabharwal A, Shah P, Button K (2022) Is depression associated with reduced optimistic belief updating?. *Royal Society Open Science* **9**:190814 https://doi.org/10.1098/rsos.190814 | PubMed

Hoogervorst K, Banellis L, Fardo F, XUE K, Rahnev D, Allen M (2024) Gender Differences in Metacognition: Global and Local Contrasts in Bias and Efficiency. *PsyArXiv* 2gwky_v1 https://doi.org/10.31234/osf.io/2gwky

Hoven M, Lebreton M, Engelmann JB, Denys D, Luigjes J, van Holst RJ (2019) Abnormalities of confidence in psychiatry: An overview and future perspectives. *Translational Psychiatry* **9**:1-18 https://doi.org/10.1038/s41398-019-0602-7 | PubMed

Hoven M, Luigjes J, Denys D, Rouault M, van Holst RJ (2023) How do confidence and self-beliefs relate in psychopathology: A transdiagnostic approach. *Nature Mental Health* **1**:337-345 https://doi.org/10.1038/s44220-023-00062-8

Katyal S, Fleming SM (2026) Gender and anxiety reveal distinct computational sources of underconfidence. *Psychological Medicine* **56**:e17 https://doi.org/10.1017/S0033291725102808 | PubMed

Katyal S, Huys QJ, Dolan RJ, Fleming SM (2025) Distorted learning from local metacognition supports transdiagnostic underconfidence. *Nature Communications* **16**:1854 https://doi.org/10.1038/s41467-025-57040-0 | PubMed

Koellinger P, Treffers T (2015) Joy Leads to Overconfidence, and a Simple Countermeasure. *PLOS One* **10**:e0143263 https://doi.org/10.1371/journal.pone.0143263 | PubMed

Korn CW, Sharot T, Walter H, Heekeren HR, Dolan RJ (2014) Depression is related to an absence of optimistically biased belief updating about future life events. *Psychological Medicine* **44**:579-592 https://doi.org/10.1017/S0033291713001074 | PubMed

Kube T, Rauch L (2025) “It’s safer to believe that others don’t like me” – A qualitative study on the paradoxical value of negative core beliefs in depression. *Behaviour Research and Therapy* **184**:104665 https://doi.org/10.1016/j.brat.2024.104665 | PubMed

Kube T, Rief W, Gollwitzer M, Gärtner T, Glombiewski JA (2019) Why dysfunctional expectations in depression persist – Results from two experimental studies investigating cognitive immunization. *Psychological Medicine* **49**:1532-1544 https://doi.org/10.1017/S0033291718002106 | PubMed

Kube T, Schwarting R, Rozenkrantz L, Glombiewski JA, Rief W (2020) Distorted Cognitive Processes in Major Depression: A Predictive Processing Perspective. *Biological Psychiatry* **87**:388-398 https://doi.org/10.1016/j.biopsych.2019.07.017 | PubMed

Lefebvre G, Lebreton M, Meyniel F, Bourgeois-Gironde S, Palminteri S (2017) Behavioural and neural characterization of optimistic reinforcement learning. *Nature Human Behaviour* **1**:0067 https://doi.org/10.1038/s41562-017-0067

Liddell TM, Kruschke JK (2018) Analyzing ordinal data with metric models: What could possibly go wrong?. *Journal of Experimental Social Psychology* **79**:328-348 https://doi.org/10.1016/j.jesp.2018.08.009

McWilliams A, Bibby H, Steinbeis N, David AS, Fleming SM (2023) Age-related decreases in global metacognition are independent of local metacognition and task performance. *Cognition* **235**:105389 https://doi.org/10.1016/j.cognition.2023.105389 | PubMed

Moses-Payne ME, Rollwage M, Fleming SM, Roiser JP (2019) Postdecision Evidence Integration and Depressive Symptoms. *Frontiers in Psychiatry* **10** https://doi.org/10.3389/fpsyt.2019.00639 | PubMed

Myin-Germeys I, Kasanova Z, Vaessen T, Vachon H, Kirtley O, Viechtbauer W, Reininghaus U (2018) Experience sampling methodology in mental health research: New insights and technical developments. *World Psychiatry* **17**:123-132 https://doi.org/10.1002/wps.20513 | PubMed

Norem JK, Cantor N (1986) Defensive pessimism: Harnessing anxiety as motivation. *Journal of Personality and Social Psychology* **51**:1208-1217 https://doi.org/10.1037/0022-3514.51.6.1208

Nüst D, Eddelbuettel D, Bennett D, Cannoodt R, Clark D, Daroczi G, Edmondson M, Fay C, Hughes E, Kjeldgaard L, *et al.* (2020) The Rockerverse: Packages and Applications for Containerization with R. *The R Journal* **12**:437 https://doi.org/10.32614/RJ-2020-007

Nüst D, Sochat V, Marwick B, Eglen SJ, Head T, Hirst T, Evans BD (2020) Ten simple rules for writing Dockerfiles for reproducible data science. *PLOS Computational Biology* **16**:e1008316 https://doi.org/10.1371/journal.pcbi.1008316 | PubMed

Rouault M, Dayan P, Fleming SM (2019) Forming global estimates of self-performance from local confidence. *Nature Communications* **10**:1141 https://doi.org/10.1038/s41467-019-09075-3 | PubMed

Rouault M, Seow T, Gillan CM, Fleming SM (2018) Psychiatric Symptom Dimensions Are Associated With Dissociable Shifts in Metacognition but Not Task Performance. *Biological Psychiatry, Translating Biology to Treatment in Schizophrenia* **84**:443-451 https://doi.org/10.1016/j.biopsych.2017.12.017 | PubMed

Rush AJ, Carmody TJ, Reimitz P.-E (2016) 16-item Quick Inventory of Depressive Symptomatology. APA PsycTests. https://doi.org/10.1037/t51259-000

Seow TXF, Fleming SM, Hauser TU (2025) Metacognitive biases in anxiety-depression and compulsivity extend across perception and memory. *PLOS Mental Health* **2**:e0000259 https://doi.org/10.1371/journal.pmen.0000259 | PubMed

Seow TXF, Gillan CM (2020) Transdiagnostic Phenotyping Reveals a Host of Metacognitive Deficits Implicated in Compulsivity. *Scientific Reports* **10**:2883 https://doi.org/10.1038/s41598-020-59646-4 | PubMed

Seow TXF, Rouault M, Gillan CM, Fleming SM (2021) How Local and Global Metacognition Shape Mental Health. *Biological Psychiatry* **90**:436-446 https://doi.org/10.1016/j.biopsych.2021.05.013 | PubMed

Sharot T, Garrett N (2016) Forming Beliefs: Why Valence Matters. *Trends in Cognitive Sciences* **20**:25-33 https://doi.org/10.1016/j.tics.2015.11.002 | PubMed

Sharot T, Korn CW, Dolan RJ (2011) How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience* **14**:1475-1479 https://doi.org/10.1038/nn.2949 | PubMed

Stefan AM, Gronau QF, Schönbrodt FD, Wagenmakers E.-J (2019) A tutorial on Bayes Factor Design Analysis using an informed prior. *Behavior Research Methods* **51**:1042-1058 https://doi.org/10.3758/s13428-018-01189-8 | PubMed

Thomas ML, Kaufmann CN, Palmer BW, Depp CA, Martin AS, Glorioso DK, Thompson WK, Jeste DV (2016) Paradoxical Trend for Improvement in Mental Health with Aging: A Community-Based Study of 1,546 Adults Aged 21–100 Years. *The Journal of Clinical Psychiatry* **77**:e1019–e1025 https://doi.org/10.4088/JCP.16m10671 | PubMed

Thompson RJ, Bailen NH, English T (2021) Everyday Emotional Experiences in Current and Remitted Major Depressive Disorder: An Experience-Sampling Study. *Clinical Psychological Science* **9**:866-878 https://doi.org/10.1177/2167702621992082

Van Marcke H, Denmat PL, Verguts T, Desender K (2024) Manipulating Prior Beliefs Causally Induces Under- and Overconfidence. *Psychological Science* **35**:358-375 https://doi.org/10.1177/09567976241231572 | PubMed

Villano WJ, Heller AS (2024) Depression is associated with blunted affective responses to naturalistic reward prediction errors. *Psychological Medicine* **54**:1956-1964 https://doi.org/10.1017/S0033291724000047 | PubMed

Villano WJ, Kraus NI, Reneau TR, Jaso BA, Otto AR, Heller AS (2023) Individual differences in naturalistic learning link negative emotionality to the development of anxiety. *Science Advances* **9**:eadd2976 https://doi.org/10.1126/sciadv.add2976 | PubMed

Xue K, Zheng Y, Papalexandrou C, Hoogervorst K, Allen M, Rahnev D (2024) No gender difference in confidence or metacognitive ability in perceptual decision making. *iScience* **27** https://doi.org/10.1016/j.isci.2024.111375 | PubMed

Zimmerman M, Ellison W, Young D, Chelminski I, Dalrymple K (2015) How many different ways do patients meet the diagnostic criteria for major depressive disorder?. *Comprehensive Psychiatry* **56**:29-34 https://doi.org/10.1016/j.comppsych.2014.09.007 | PubMed

Peer reviews

Reviewer #1 (Public review):

Summary:

The authors investigate how depressive symptoms relate to metacognitive confidence across multiple levels of a metacognitive hierarchy. Participants completed twice-daily assessments of depressive symptoms and bi-daily assessments of a perceptual confidence task for eight weeks. The study replicates prior findings that depression is associated with lower confidence and extends these findings by suggesting that trait depression weakens the temporal persistence of local confidence signals, thereby impairing their accumulation into global confidence.

The study addresses an important question in computational psychiatry: how disturbances in confidence contribute to persistent negative self-beliefs in depression. The longitudinal design and repeated assessments represent advances over prior cross-sectional studies. The attempt to bridge momentary confidence fluctuations with broader self-beliefs through a hierarchical metacognitive framework is novel.

The principal contribution is the finding that trait depression moderates the lagged relationship between local and global confidence. The authors interpret this as evidence that positive fluctuations in local confidence decay more rapidly in individuals with higher depression, limiting their integration into higher-order confidence beliefs. This contribution is potentially important.

However, several aspects of the interpretation warrant caution. The moderated mediation analysis remains correlational and does not establish that impaired local confidence persistence causally produces global under-confidence. Alternative explanations, including stable individual differences in response styles, latent third variables, or measurement properties of the confidence scales, remain plausible. The manuscript occasionally adopts language suggesting mechanistic or causal conclusions that exceed the inferential scope of the analyses.

The temporal resolution of the study also complicates interpretation. The absence of cross-lagged effects between depression and confidence may reflect a mismatch between the timescales over which mood and metacognition interact. The authors acknowledge this possibility, but it substantially limits conclusions regarding temporal precedence.

Furthermore, the sample size is not justified, and the sample differs considerably from populations typically studied in depression research. Participants were older, predominantly female, and self-selected citizen scientists with low and relatively stable depression scores. Consequently, it remains unclear whether the observed dynamics generalise to clinically depressed populations, where symptom severity and variability may differ substantially.

Overall, this is a thoughtful and technically sophisticated study that provides valuable new data on the temporal organisation of confidence in relation to depression. The central findings are interesting and likely to stimulate future work, although the mechanistic interpretations would benefit from greater caution.

Strengths:

(1) Innovative use of dense longitudinal sampling to investigate metacognitive processes.

(2) Large number of repeated observations per participant and good adherence over eight weeks.

(3) Integration of EMA, multilevel vector autoregression, Bayesian modelling, and computational modelling.

(4) Novel proposal that depression weakens the persistence of local confidence signals and their integration into global confidence.

(5) Careful consideration of local versus global metacognitive processes.

Weaknesses:

(1) The causal and mechanistic claims may exceed what can be inferred from the data.

(2) No justification is provided for the sample size, and the sample is older, predominantly female, and largely non-clinical, limiting generalisability.

(3) The sampling intervals may not be optimally suited to detect temporal relationships between mood and confidence.

(4) Several modelling decisions require additional justification and sensitivity analyses.

(5) The moderated mediation framework assumes a temporal ordering that cannot be conclusively established.

https://doi.org/10.7554/eLife.112173.1.sa2

Reviewer #2 (Public review):

Summary:

Phon-Amnuaisuk and colleagues address an important question in cognitive psychology and computational psychiatry: how does the established relationship between depression and metacognitive under-confidence unfold over time? The study is grounded in a hierarchical view of metacognition, in which local confidence in individual decisions contributes to global estimates of performance. To test this, the authors used an intensive longitudinal design in which participants repeatedly reported depressive symptoms and completed a gamified perceptual decision-making task over eight weeks. This allowed them to examine whether depressive symptoms and confidence fluctuate together within individuals, whether one predicts the other over time, and whether trait depression alters the way local confidence is carried forward and integrated into global confidence. The main findings are that higher trait depression is associated with lower local and global confidence, that within-person mood fluctuations show limited temporal precedence over confidence at the two-day timescale, and that trait depression is linked to weaker temporal persistence of local confidence and reduced carry-over into later global confidence.

Strengths:

A major strength of the study is its repeated-measures design across a large sample. Participants completed up to 28 metacognitive task sessions over eight weeks, alongside repeated ratings of depressive symptoms. This allows the authors to separate stable between-person differences from within-person changes over time, which is a clear advantage over standard cross-sectional studies. The analytical strategy is also appropriate: the authors use multilevel vector autoregressive models to examine temporal, contemporaneous, and between-person associations; Bayesian ordinal models to test interactions; and computational modelling to examine how confidence is formed.

The strongest results concern stable individual differences. Participants with higher average depressive symptoms reported lower local and global confidence. While this pattern is consistent with prior work showing reduced confidence in depression, the computational model further suggests that higher depression is associated with a more conservative confidence criterion: these participants required more evidence before reporting high confidence. This adds nuance by suggesting that under-confidence in depression may not reflect poorer metacognitive sensitivity, but rather a bias in how confidence is reported.

The study also proposes a novel temporal account. Trait depression was associated with weaker autocorrelation of local confidence across sessions, which in turn reduced the extent to which local confidence carried over into later global confidence. This indirect pathway was supported by a moderated mediation analysis and was consistent across most individual depressive symptoms. The finding is theoretically interesting and supported by strong statistical evidence.

Weaknesses:

The main limitation concerns the interpretation of the temporal and mechanistic claims. The strongest effects are observed at the trait level, whereas the within-person temporal associations between mood and confidence are weak or absent. The study therefore provides stronger evidence that people with higher average depressive symptoms are generally less confident than evidence that momentary changes in depressive mood drive later changes in confidence, or vice versa.

The sample also constrains the conclusions. Depression scores were strongly concentrated near the lower end of the scale, and the final sample was highly selected, with many of the initial 976 participants excluded because they did not complete enough task sessions. This means that the study may be better suited to detecting stable individual differences than dynamic mood-confidence processes.

The temporal spacing of the metacognitive assessments is a further constraint. Because the metacognition task was administered every two days, the cross-lagged analyses could only test mood-confidence dynamics across this interval. If depressive mood and confidence influence each other over shorter timescales, such effects may have been missed. The absence of cross-lagged effects should therefore be interpreted with caution: it shows that temporal precedence was not detected at the two-day lag in this sample, but it does not rule out shorter-term directional effects or effects in more symptomatic clinical populations.

Related to this, the manuscript moves between several related but distinct terms - "mood," "depressive mood," "depression," "depressive symptoms," and "trait depression" - without always clarifying whether these are intended as interchangeable or as conceptually distinct constructs. This matters for a study whose central claims concern temporal precedence and trait-versus-state distinctions, and it is compounded by a sample with generally low depressive symptom levels, where the boundary between transient low mood and a trait-like depressive disposition is harder to draw.

https://doi.org/10.7554/eLife.112173.1.sa1

Author response:

Reviewer #1:

We thank the reviewer for their comments. They raised an issue with the correlational nature of the analysis, suggesting alternative explanations, including stable individual differences in response styles, latent third variables, or measurement properties of the confidence scales. We agree with the reviewer’s comments that a latent third variable may confound our findings and higher-order metacognitive beliefs (which we did not assess) are a possible candidate. We will broaden our discussion to include other plausible confounds that may jointly relate to depression and confidence dynamics. Regarding stable individual differences in response styles, our analysis decomposed confidence into within and between-person components and standardised the within-person component by each participant’s own variability. This allowed the interaction and moderated mediation analyses to separate within-person associations from stable between-person differences (e.g., range of confidence scale used; Epskamp et al., 2018). With regards to measurement properties, we will conduct additional analyses investigating whether trait depression is associated with altered response patterns (e.g., non-linear or more variable mapping of latent evidence into confidence reports).

The reviewer also noted that the temporal resolution we selected may not be optimal to measure co-fluctuation of depression and confidence. We agree this remains a challenge for the depression research (Jamalabadi et al., 2025; Tamm et al., 2024), and metacognitive confidence research (da Fonseca et al., 2023; Wright et al., 2024). We chose the interval as it allowed us to balance retention and data quality across an 8-week study (Eisele et al., 2022) and address a gap in the literature – frequent but shorter (~1-week) EMA-based assessments of existing metacognitive confidence studies (da Fonseca et al., 2023; Wright et al., 2024) precludes insights into the dynamics of mood and metacognitive confidence across longer timescales. We have explored one-day and half-day lagged associations but did not find significant effects across most symptoms for both local and global confidence. This was omitted from the original manuscript as the analyses could not symmetrically control for local/global confidence (as these were only assessed bi-daily) but will be included in the revised manuscript's supplement.

The reviewer noted that the temporal ordering of the moderated mediation analysis cannot be conclusively established. We agree that there is a lack of research investigating alternative orderings. We selected the local-to-global ordering because prior work has similarly modelled and demonstrated global confidence estimates as a cumulative integration of local confidence estimates across the block (Cavalan et al., 2023; Katyal et al., 2025; Lee et al., 2021; Rouault et al., 2019, 2022). Nevertheless, we modelled an alternative process (i.e. whether global confidence interacted with trait-level depression in predicting subsequent local confidence) but did not find a significant frequentist interaction effect. We will elaborate on our rationale for the current ordering and include analyses of this alternative process in our revised manuscript and supplement.

Finally, we agree with the reviewer's comments that the sample constraints generalisability. The revised manuscript will more clearly discuss the generalisability of our findings to clinical populations and the implications of self-selection and limited within-person variability in depressive symptoms for interpretation and future work. We will also tone down the causal language in the manuscript.

Reviewer #2:

We thank the reviewer for their comments. We agree with the reviewer that our strongest effects live at the trait level, whereas our cross-lagged findings provided insufficient evidence for depression driving underconfidence or vice versa (at least with a two-day lag). However, our moderated mediation finding centres on a different angle - trait level depression moderating how within-person confidence is integrated into more global beliefs. Nevertheless, we agree that the findings provide a plausible account for why underconfidence (and possibly metacognitive beliefs) remain persistent in depression, but not how either underconfidence or depression arose in the first place. We will make this clearer in our revised manuscript.

Consistent with reviewer #1's comments, we agree that the non-clinical nature of sample does constrain generalisability and inference. We will conduct a sensitivity analysis with a larger sample (more lenient inclusion criteria) and discuss its implications in more detail in our revised manuscript. We also agree with the reviewer's comments that our findings do not preclude the possibility of temporal precedence and will further clarify in our revised manuscript with reference to shorter time lags. The reviewer noted that our use of terminologies (e.g., "depression", "depressive symptoms", "depressive mood") was not clearly defined. Our revised manuscript will make this point clearer whilst also making the usage of terminologies consistent.

References

Cavalan, Q., Vergnaud, J.-C., & de Gardelle, V. (2023). From local to global estimations of confidence in perceptual decisions. Journal of Experimental Psychology: General, 152(9), 2544–2558. *https://doi.org/10.1037/xge0001411* da Fonseca, M., Maffei, G., Moreno-Bote, R., & Hyafil, A. (2023). Mood and implicit confidence independently fluctuate at different time scales. Cognitive, Affective, & Behavioral Neuroscience, 23(1), 142–161. *https://doi.org/10.3758/s13415-022-01038-4*

Eisele, G., Vachon, H., Lafit, G., Kuppens, P., Houben, M., Myin-Germeys, I., & Viechtbauer, W. (2022). The Effects of Sampling Frequency and Questionnaire Length on Perceived Burden, Compliance, and Careless Responding in Experience Sampling Data in a Student Population. Assessment, 29(2), 136–151. *https://doi.org/10.1177/1073191120957102*

Epskamp, S., Waldorp, L. J., Mõttus, R., & Borsboom, D. (2018). The Gaussian Graphical Model in Cross-Sectional and Time-Series Data. Multivariate Behavioral Research, 53(4), 453–480. *https://doi.org/10.1080/00273171.2018.1454823*

Jamalabadi, H., Koosha, T. A., Stocker, E., Jansen, A., Ebner-Priemer, U. W., Proppert, R. K. K., Rieble, C. L., Tutunji, R., & Fried, E. I. (2025). Optimizing the frequency of ecological momentary assessments using signal processing. Psychological Medicine, 55, e358. *https://doi.org/10.1017/S003329172510264X*

Katyal, S., Huys, Q. J., Dolan, R. J., & Fleming, S. M. (2025). Distorted learning from local metacognition supports transdiagnostic underconfidence. Nature Communications, 16(1), 1854. *https://doi.org/10.1038/s41467-025-57040-0*

Lee, A. L. F., de Gardelle, V., & Mamassian, P. (2021). Global visual confidence. Psychonomic Bulletin & Review, 28(4), 1233–1242. *https://doi.org/10.3758/s13423-020-01869-7*

Rouault, M., Dayan, P., & Fleming, S. M. (2019). Forming global estimates of self-performance from local confidence. Nature Communications, 10(1), 1141. *https://doi.org/10.1038/s41467-019-09075-3*

Rouault, M., Will, G.-J., Fleming, S. M., & Dolan, R. J. (2022). Low self-esteem and the formation of global self-performance estimates in emerging adulthood. Translational Psychiatry, 12(1), 1–10. *https://doi.org/10.1038/s41398-022-02031-8*

Tamm, J., Takano, K., Just, L., Ehring, T., Rosenkranz, T., & Kopf-Beck, J. (2024). Ecological Momentary Assessment versus Weekly Questionnaire Assessment of Change in Depression. Depression and Anxiety, 2024, 9191823. *https://doi.org/10.1155/2024/9191823*

Wright, A. C., Palmer-Cooper, E., Cella, M., McGuire, N., Montagnese, M., Dlugunovych, V., Liu, C.-W. J., Wykes, T., & Cather, C. (2024). Experiencing hallucinations in daily life: The role of metacognition. Schizophrenia Research, Hallucinations: Neurobiology and Patient Experience, 265, 74–82. *https://doi.org/10.1016/j.schres.2022.12.023*

https://doi.org/10.7554/eLife.112173.1.sa0