New Ideas in Psychology 74 (2024) 101092

Contents lists available at ScienceDirect

New Ideas in Psychology

journal homepage: www.elsevier.com/locate/newideapsych

An active inference approach to interpersonal differences in depression

Lukas Kirchner[a,*,1], Anna-Lena Eckert[b,1], Max Berg[a], Dominik Endres[b], Benjamin Straube[c], Winfried Rief[a]

[a] Department of Psychology, Clinical Psychology and Psychotherapy, Philipps-University of Marburg, Germany
[b] Department of Psychology, Theoretical Cognitive Science, Philipps-University of Marburg, Germany
[c] Department of Psychiatry and Psychotherapy, Philipps-University of Marburg, Germany

ARTICLE INFO

Keywords:
depression
active inference
predictive processing
social expectations
interpersonal beliefs
social behavior
interpersonal functioning

ABSTRACT

Depression is characterized by different distortions in interpersonal experience and behavior, ranging from social withdrawal to overt hostility. However, clinical psychological research has largely neglected the need for an integrative framework to operationalize these different phenomena and their dynamic change more accurately in depression. In this article, we draw on active inference theory, a comprehensive theory of perception, action, and learning, to provide a formal model explaining how variations in patients' internal belief-systems lead to differences in social experience and behavior. In this context, we assume that individuals cannot directly grasp the characteristics of their social environment. Instead, they must infer them indirectly from ambiguous social observations, which they themselves generate and alter through their actions. Differences in interpersonal experience and behavior arise from the interplay of patients' prior expectations, their propensity to infer particular social states from certain observations, and their beliefs in their ability to influence these situations through specific actions. We then use concrete examples to demonstrate how future research can take our approach to identify systematic differences in interpersonal experiences and behaviors among depressed patients (or patient subgroups) and to investigate their changes in response to new social experiences. We also discuss potential applications of our approach in diagnosing and treating depression. This work is a move towards understanding the interpersonal aspects of depression in more detail, recognizing their importance in etiology, diagnosis, and treatment.

1. Introduction

Since Coyne's (1976) influential characterization of the interpersonal nature of depression, there has been extensive research on the interpersonal factors contributing to the development and maintenance of depressive symptoms. It is widely agreed that depressed individuals exhibit significant distortions in the anticipation, perception, and processing of interpersonal information, as well as in their interpersonal behavior (Hames, Hagan, & Joiner, 2013; Kupferberg & Hasler, 2023; Segrin, 2000). For instance, depressed individuals are more sensitive to rejection (Gao, Assink, Cipriani, & Lin, 2017) and perform worse on tasks involving social cognition (Nestor, Sutherland, & Garber, 2022). They also have difficulty with social problem-solving while interacting with others (Hames et al., 2013).

From an etiological perspective, research suggests a bidirectional relationship between depressive symptoms and impaired social functioning (Beeson, Brittain, & Vaillancourt, 2020; Kirchner, Schummer, Krug, Kube, & Rief, 2022). Negative interpersonal experiences such as peer rejection, childhood maltreatment, and lack of social support are common risk factors (Gariépy, Honkaniemi, & Quesnel-Vallée, 2016; Nanni, Uher, & Danese, 2012; Platt, Kadosh, & Lau, 2013). According to this view, interpersonal trauma and the lack of social support may trigger socially maladapted perception (Gadassi & Rafaeli, 2015), informational processing (Everaert, Podina, & Koster, 2017), and behavior (Evraire & Dozois, 2011), which in turn promotes new negative interpersonal experiences (Starr & Davila, 2008).[2] On the other hand, depressive symptoms might themselves impair social skills in the long run, leading also to new negative interpersonal experiences

* Corresponding author. Department of Clinical Psychology and Psychotherapy, University of Marburg, Gutenbergstraße 18, 35037, Marburg, Germany.
E-mail address: lukas.kirchner@uni-marburg.de (L. Kirchner).

[1] Both authors contributed equally to this work.

[2] This etiological model is especially relevant to chronic forms of depression, where interpersonal trauma and interpersonal problems are core features (Schramm et al., 2020).

https://doi.org/10.1016/j.newideapsych.2024.101092
Received 17 July 2023; Received in revised form 11 March 2024; Accepted 14 April 2024
Available online 26 April 2024
0732-118X/

(Dobson, Quigley, & Dozois, 2014).

Depression has been associated with a variety of interpersonal distortions (Bird, Tarsia, & Schwannauer, 2018; Bourke, Douglas, & Porter, 2010; Dobson et al., 2014; Gadassi & Rafaeli, 2015; Hames et al., 2013; Hirschfeld et al., 2000; Kupferberg, Bicks, & Hasler, 2016; Kupferberg & Hasler, 2023; Liu & Alloy, 2010; Nestor et al., 2022; Segrin, 1990, 2000; Starr & Davila, 2008; Teichman & Teichman, 1990; Tse & Bond, 2004; Weightman, Air, & Baune, 2014). However, traditional research has largely neglected the precise operationalization of these different manifestations of depression (for laudable exceptions, see, Cain et al., 2012; Dawood, Thomas, Wright, & Hopwood, 2013; Simon, Cain, Wallner Samstag, Meehan, & Muran, 2015), which extends beyond the interpersonal symptom domain (de Jonge, Wardenaar, & Wichers, 2015; Fried, Flake, & Robinaugh, 2022; Goldberg, 2011; Monroe & Anderson, 2015).

For instance, depression can be associated with both inhibited interpersonal behaviors such as social withdrawal, and disinhibited behaviors like excessive reassurance-seeking (Hames et al., 2013; Kupferberg & Hasler, 2023). Moreover, while depression often manifests in terms of dependent, submissive, and anxiously-avoidant experiences and behaviors, it can also involve impulsive and hostile behaviors including retaliation or uncontrolled expressions of anger (Bird et al., 2018; Hames et al., 2013; Kupferberg & Hasler, 2023). Clinicians have long recognized the varied interpersonal manifestations of depression, and tailor their treatments accordingly. Intervention strategies include psychoeducation (Cuijpers, Muñoz, Clarke, & Lewinsohn, 2009), interpersonal context discrimination training (McCullough, Schramm, & Penberthy, 2014; McCullough, 2006a), behavioral activation (Uphoff et al., 2020), experiential and emotion-focused techniques (Renner, Arntz, Leeuw, & Huibers, 2013; Roediger, Stevens, & Brockman, 2018), role disputes (Klerman & Weissman, 1994), training of new interpersonal skills (Bellack, Hersen, & Himmelhoch, 1996; Lipsitz & Markowitz, 2013), repeated social exposure (Muñoz, Ghosh Ippen, Rao, Le, & Valdes Dwyer, 2000), disciplined personal involvement (McCullough, 2006b), and behavioral experiments (Bennett-Levy et al., 2004). A more precise operationalization of such diverse interpersonal manifestations of depression would offer opportunities to better understand etiological, diagnostic, and treatment-related differences in depression.

Furthermore, traditional research approaches have had limited success in explaining and predicting how depressed individuals' interpersonal experiences and behaviors adapt dynamically in response to changing interpersonal environments (Wichers, 2014). For example, there is evidence that depressed individuals struggle to adjust negative interpersonal beliefs in response to positive interpersonal experiences (Kube, 2023), favoring experiences that align with their existing beliefs (e.g., "As I said, no one talked to me at the party!") (Badcock, Davey, Whittle, Allen, & Friston, 2017).[3] A more precise (i.e., mathematical) description and examination of how interpersonal experiences are integrated into existing belief-systems and how these belief-systems change dynamically over time in response to new social experiences would provide an opportunity for more accurate understanding of how biased belief-systems and pathological interpersonal behaviors in depression can be altered progressively through treatment.

Relevance of computational models. Computational models, with their mathematical formalism and flexibility, offer a precise means of describing heterogeneous experiences and behaviors, and of modelling dynamic adaptations in response to environmental changes (Friston et al., 2016; Montague, Dolan, Friston, & Dayan, 2012; Schwartenbeck & Friston, 2016). They provide a unified framework to formalize hypotheses, functionally link and simulate internal experiences and external behaviors, and derive clear-cut predictions about how these experiences and behaviors change in response to observations from the external environment (Hauser, Skvortsova, De Choudhury, & Koutsouleris, 2022). Accordingly, computational approaches can enable researchers to operationalize and explore heterogeneity in interpersonal manifestation of depression more accurately and systematically. Moreover, they can help researchers to better understand under which conditions these symptoms develop, manifest, persist, and change in response to the social environment (Smith, Badcock, & Friston, 2021; Story et al., 2023).

Aims of this article. In summary, research suggests that distortions in interpersonal experience and behavior are prevalent and varied in depression. However, there is still a lack of formal approaches to model this heterogeneity and the dynamic change in interpersonal experience and behavior in response to the social environment.

Compared to traditional methods, formalizing interpersonal processes within an integrative computational framework enables more precise operationalization of different interpersonal phenomena in depression. It also aids in understanding the environmental conditions that influence how such experiences and behaviors develop, persist, and change over time.

This article applies the theory of active inference, a comprehensive theory of perception, action, and learning, to provide such a formalization. Our approach functionally connects distorted interpersonal experiences in depression, such as vague or negative social expectations, low confidence in social observations, negative interpretation of these observations, and low efficacy expectations about one's own behavior, with distorted interpersonal behavior such as dependent, hostile, passive, or avoidant means of action. Furthermore, it enables a systematic investigation of how these processes dynamically change in response to the social environment.

In the subsequent sections, we focus on the structure, motivation, and functional form of the model, rather than providing numerical analyses using a specific example. In the final section, we discuss potential applications and ideas for implementation.

2. Modelling interpersonal differences in depression

Introduction to active inference. In recent decades, emerging computational theories have revolutionized our understanding of the mind in both health and psychological disorders. In cognitive and computational neuroscience, the active inference framework enables a unified approach to complex processes such as perception, learning, action, and decision-making via the Bayesian decision theory (Da Costa et al., 2020; Huys, Guitart-Masip, Dolan, & Dayan, 2015). At the core of active inference is the concept of the mind as an intricate inference engine, constantly trying to make sense of its environment based on noisy, ambiguous, and incomplete sensory information (Friston, FitzGerald, Rigoli, Schwartenbeck, & Pezzulo, 2017). The external world's "reality" is not directly accessible to the agent; instead, the agent must infer the "hidden states" of the world using its internal "generative model" (Heins et al., 2022; Smith, Friston, & Whyte, 2022). This generative model represents the agent's beliefs about how sensory information is related to environmental states, and how states lead to observable consequences. The agent's probabilistic model of the world is continuously updated with sensory input to ensure the accurate

[3] Explanations from general psychology and clinical psychology suggest that several mechanisms are involved in this phenomenon. These include learned helplessness (Seligman, 1972), neurobiological abnormalities (Ng, Alloy, & Smith, 2019; Porcelli et al., 2019; Yang et al., 2022), increased neural effort when processing social cues (Suffel et al., 2020), pre-operational thinking (J. P. McCullough, 2003), dysfunctional core beliefs (Beck, 1963, 1964), negative relational schemas (Baldwin, 1992), blunted elaboration of positive information (Chen, Takahashi, Nakagawa, Inoue, & Kusumi, 2015; Halahakoon et al., 2020), social anhedonia (Barkus & Badcock, 2019), persistent negative expectations (Kube et al., 2020; Roepke & Seligman, 2016), and cognitive immunization (Rief & Joormann, 2019). In addition to these findings, evidence suggests that depressed individuals have trouble drawing accurate conclusions from social interactions (Weightman et al., 2014), and demonstrate abnormal patterns of reward-learning and decision-making in social contexts (Safra, Chevallier, & Palminteri, 2019).

inference of hidden states. It is combined with the agent's "prior" beliefs about specific hidden states or observations (Friston et al., 2016; Smith, Badcock, & Friston, 2021). Having accurate prior beliefs and generative models is essential for the agent's optimal adaptation to its environment (Da Costa et al., 2020).

A crucial aspect of active inference is action. By performing action sequences, or "policies," the agent can influence the environment's hidden states (Heins et al., 2022). Inferring policies allows the agent to select those action sequences that result in desired or preferred outcomes or sensory information. In other words, action helps to minimize surprise in the long term.[4] This fundamental understanding of sentient behavior involves choosing action plans that produce predicted consequences (Adams, Shipp, & Friston, 2013). When applied to social behaviors and decision-making, this principle is often referred to as "planning as inference" (Attias, 2003; Botvinick & Toussaint, 2012; Kaplan & Friston, 2018). In other words, policy selection involves committing to actions that minimize deviations from predicted or expected outcomes.

From an evolutionary perspective, active inference involves constantly making a trade-off between seeking new information through action, and maximizing access to known rewards to ensure survive in an ever-changing environment (Gopnik, 2020; Kouvaris, Clune, Kounios, Brede, & Watson, 2015). Agents who fail to engage in any exploratory behavior and only seek to maximize known rewards "over-fit" their environment, putting themselves at risk of being incapable of adapting to environmental changes. On the other hand, "under-fitting" agents who engage in a lot of exploratory behavior and show little reward maximization run the risk of failing to exploit resources and rewards to maintain their fitness.

Active inference suggests that mental disorders such as depression result from a biased inference of hidden states in the world. This bias is attributable to a biased generative model (Schwartenbeck et al., 2015). Bayesian formulations have been useful in this context in explaining depression's different symptom domains (Berg, Feldmann, Kirchner, & Kube, 2022; Chekroud, 2015; Smith, Alkozei, Killgore, & Lane, 2018). For example, Smith et al. (2018) proposed that depressive schemas[5] may serve as precise, internalized priors in individuals with depression. They also suggest that psychomotor slowing and loss of energy, key symptoms of depression, may arise from biased beliefs about the effectiveness of actions in achieving desired sensory states. From this perspective, the mind may infer that the cost of certain actions "outweighs" their utility, thus lowering motivation for action and encouraging the tendency for depressed individuals to stay in bed (Smith et al., 2018). These costs may be high particularly when processing social information because positive social cues are less expected (Suffel, Nagels, Steines, Kircher, & Straube, 2020).

Applying active inference theory to interpersonal phenomena in depression. Human social interactions are incredibly complex. During everyday conversations, individuals must consider a multitude of factors, including noisy verbal and non-verbal information, as well as ambiguous cultural, contextual, and situational elements. These factors evolve constantly and are sometimes unpredictable. From a Bayesian perspective, individuals must process highly uncertain information to infer the hidden characteristics of their interpersonal environment and respond in real time to gain interpersonal information and meet interpersonal needs.

Individuals with depression struggle to draw accurate conclusions from social interactions (Weightman et al., 2014). Moreover, they exhibit different distortions in interpersonal experience and behavior that can vary among (but also within) affected individuals (Hames et al., 2013; Kupferberg & Hasler, 2023).

We propose that the common and sometimes even contradictory interpersonal behaviors seen in depression, such as excessive reassurance-seeking, retaliation, passivity, or avoidance (Hames et al., 2013; Kupferberg & Hasler, 2023), result from the interplay of different biases in inferring the hidden characteristics of the interpersonal environment through observation and action.

The active inference approach, with its inherent flexibility, allows us to represent these different biases within depressed individuals' generative models which integrate their beliefs about how sensory information relates to specific characteristics in the social environment and their actions within it.[6] For example, depressed individuals' generative models can reflect vague or negative expectations regarding future events, having low confidence in one's own observations or interpreting such observations in a negative manner, and holding low self-efficacy expectations in general or high efficacy-expectations about dysfunctional action sequences (Everaert, 2021; Everaert et al., 2017; Gamble, Moreau, Tippett, & Addis, 2019; Kavanagh, 2014; LeMoult & Gotlib, 2019; Rief & Joormann, 2019) (Fig. 1).

Consider the scenario where someone attends a party where they only know a few people. A person without depression would likely approach the party anticipating that most guests will be friendly and engage in conversation once introduced (positive expectation regarding future events). They would interpret a smile from another table as a sign of friendliness (interpretation in a positive manner) and expect that initiating a conversation would lead to a positive social experience (high efficacy-expectations for functional interpersonal action sequences). Conversely, a person with depression might approach the party with more vague or even negative expectations about how they will be treated (vague or negative expectations regarding future events). They might question whether they truly saw a smile from the other table (low confidence in one's own observations), or may interpret it as a clear sign of pity or even mockery (interpretation in a negative manner). Their reaction, whether to withdraw (e.g., leave the party) or start an argument, would depend on how much control they believe they have over the situation (low self-efficacy expectations) and what behavior they consider to be effective (high efficacy-expectations for dysfunctional action sequences).

[4] Mathematically speaking, we can approximate the minimization of "surprise" by minimizing a quantity called "variational free energy" (VFE). In the context of active inference, minimizing VFE is expected to lead to "optimal" perceptions and actions (Friston, 2010). More generally, actions or policies are selected to minimize "expected free energy" (EFE) (Da Costa, Parr, Sengupta, & Friston, 2021). Technically, EFE encompasses two types of optimality. The first reflects the inclination to bring about desired outcomes, which are the "states of being" that characterize the decision maker or person in question. Mathematically, this inclination can be interpreted as maximizing expected value, or the (log) probability of being in a characteristic state. The other part of EFE supports information-seeking and the resolution of uncertainty. In short, active inference involves a joint commitment to optimal Bayesian decisions by maximizing expected value (Berger, 2011), while simultaneously adhering to the principles of optimum Bayesian design by maximizing expected information gain (Lindley, 1956; MacKay, 1992). Active inference effectively helps to solve the exploration-exploitation dilemma by balancing epistemic and pragmatic imperatives in well-adjusted agents (Schwartenbeck et al., 2019).

[5] Here, schemas are defined as "(...) maladaptive, pessimistic sets of beliefs/expectations that bias perceptual/conceptual inter- pretations of new sensory input, as well as the subsequent predic- tions, judgments, and decisions these interpretations inform (...)" (Smith et al., 2018, p. 375).

[6] In fact, any social experience and behavior can be seen as deviations from the probability distributions making up the generative model. This concept is known as the complete class theorem (Brown, 1981; Wald, 1947), which states that there are prior beliefs that make the behavior Bayes optimal for any combination of behaviors and value functions. In other words, the ideal Bayesian assumptions in active inference (refer to the above) allow us to describe any behavior based on an agent's prior beliefs. This is one of the main reasons for computational phenotyping, which characterizes a patient's choice behavior by examining their underlying prior beliefs (Smith, Kirlic, et al., 2021).

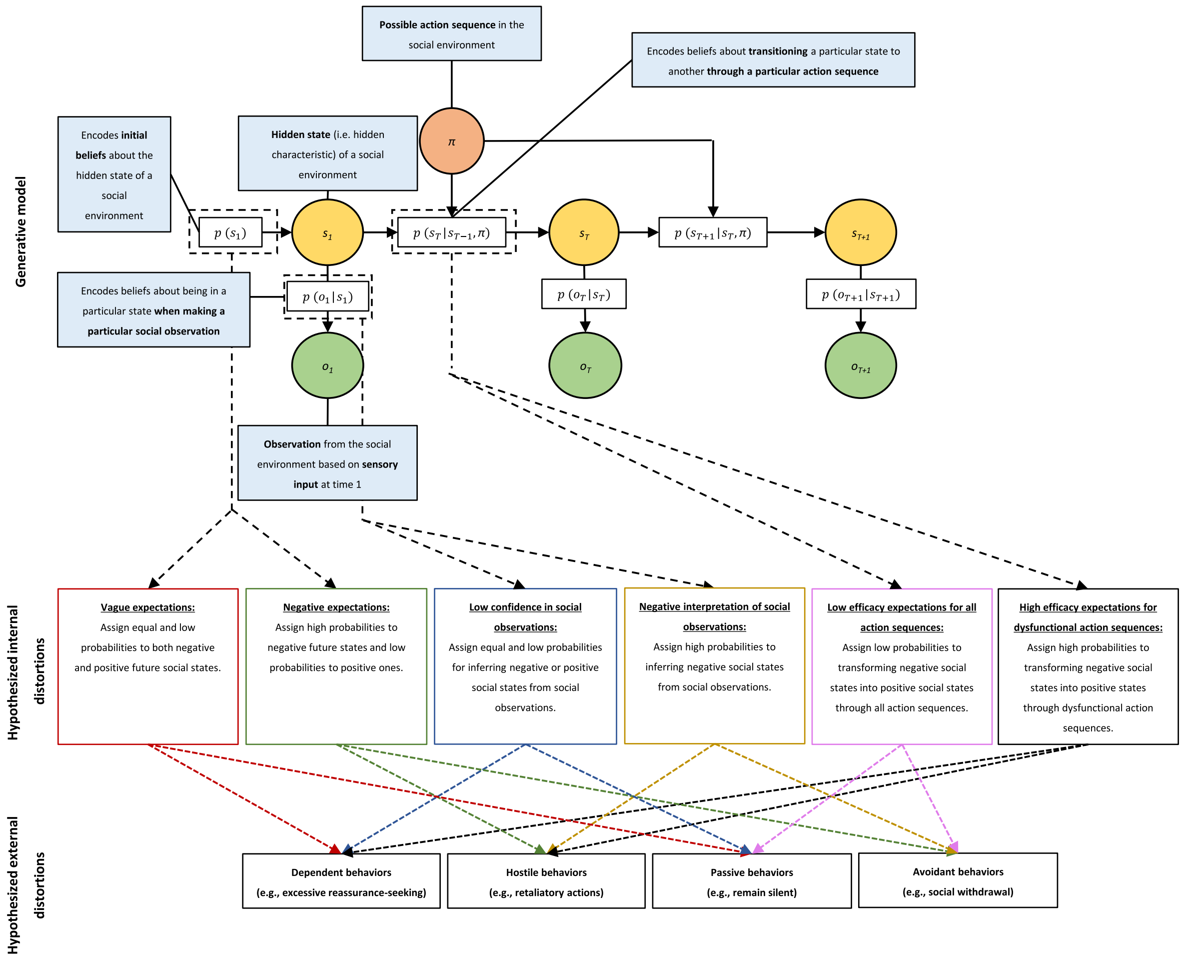

Fig. 1. Modeling active inference of hidden social states in depressed individuals as a partially observable Markov decision process. T = time, p = probability, s = hidden state of a social environment, o = observation based on sensory input from the social environment, π = possible action sequence in a social environment. Arrows indicate temporal precedence. Dashed arrows indicate hypothesized distortions in patients' internal belief-systems represented by different parameter specifications in the generative model. Dotted arrows indicate hypothesized distortions in patients' external behaviors as a function of distortions in the generative model. Note that further internal and external distortions could possibly be represented by different parameter specifications and their combination in the generative model.

Frequent social withdrawal or open hostility might further bias depressed individuals' interpersonal belief-systems in the long run, as such behaviors limit the acquisition of new (i.e., corrective) sensory input or even increase the number of negative interpersonal responses from the social environment (Kube, 2023; Kube, Schwarting, Rozenkrantz, Glombiewski, & Rief, 2020). As a consequence, the interpersonal experience and behavior of depressed individuals become increasingly disconnected from the social environment (Hames et al., 2013; Kupferberg & Hasler, 2023; Segrin, 2000). A systematic review by Nagy and Moore (2017) found that out of 24 studies employing interventions facilitating social interaction (i.e., initiating "sampling" of new social information), 17 reduced levels of depressive symptoms successfully.

Modelling biased social state inference in depression. Actively inferring the hidden states of a social environment can be understood as making decisions about actions based on sensory input while facing uncertainty about the meaning of that input and the actions' future outcomes (e.g., Moutoussis, Trujillo-Barreto, El-Deredy, Dolan, & Friston, 2014). In this context, actions have both explorative (minimizing uncertainty) and exploitative (maximizing rewards) functions (Friston et al., 2016). In active inference, a common model of how agents internally represent aspects of the external world are partially observable Markov decision process models (POMDPs; Da Costa et al., 2020). These models have been described in clinical applications (Moutoussis et al., 2014; Smith et al., 2020, 2022). POMDPs involve making decisions under uncertainty, where the outcomes depend on the environment's hidden states and are associated with rewards or costs, making them suitable for applications in active inference (Smith et al., 2022).

In this paragraph, we briefly describe the key concepts and parameters of a POMDP generative model and apply them to the active inference of social environment characteristics during brief social interactions (Fig. 1). Agents have beliefs about the state of an interpersonal context (s) and their possible action sequences in that context (π). They make observations (o) mediated by their senses. Agents also have beliefs about the probability of being in a particular interpersonal state (s_T) when making a specific observation (o_T). This is represented by the agent's observational model, which is a probability distribution $p(o_T|s_T)$ providing information about the likelihood of certain observations given specific states. Furthermore, agents entertain beliefs about the probability of transitioning from one state to another that are potentially influenced by their own interpersonal actions. This is represented by a probability distribution $p(s_{T+1}|s_T, \pi)$. Finally, agents may possess initial beliefs about the hidden state of a social environment $p(s_1)$.

In this model, deviations in the specified probability distributions, or their combinations, can explain the different interpersonal phenomena associated with depression mentioned above. For instance, we can represent vague or negative expectations about which social state will be entered as an initial prior probability ($p(s_1)$). Here, a vague or negative expectation can be represented by assigning equally low probabilities to negative and positive social states or by assigning a higher probability to a negative social state.

Individuals with depression might also display biased beliefs to what extent a certain observation reflects a certain state of the social environment. For example, they might generally have low confidence in their social observations, or infer predominantly negative social states from them. This can be represented through the patient's observation model $p(o_T|s_T)$. Here, having low confidence in a social observation can be represented by assigning equally low probabilities to infer a negative and positive social state from a social observation. The tendency to infer a negative social state from a social observation can be represented by assigning a higher probability to infer such a state from a social observation than by inferring a positive one.

Moreover, individuals with depression may have different beliefs about how possible action sequences will influence a hidden state in the social environment. For instance, they might have low expectations of changing these characteristics through any action sequence (low self-efficacy) or high expectations of changing them through dysfunctional action sequences, such as behaving in a dependent, hostile, passive, or avoidant manner. This can be represented by assigning equally low probability for transforming a negative social state to a positive one by any action sequence, or by assigning a high probability for achieving this transformation through a dysfunctional action sequence (concerns $p(s_{T+1}|s_{T,}\ \pi)$).

3. Implications and future research

Compared to traditional depression research, our approach enables a more flexible and accurate description of the diverse interpersonal phenomena in depression, and their dynamic changes in response to the social environment. By functionally linking interpersonal perception, decision-making, action, and learning within an integrative model, researchers can gain a deeper understanding of *why* depressed individuals differ in their social experience and behavior and *how* their internal belief-systems adapt in response to new interpersonal experiences.

This has potential applications in etiology and psychopathology research. For instance, it can be used to identify systematic differences between patients or patient subgroups suffering from depression (cf. Schwartenbeck & Friston, 2016). It can also help us investigate how new social experiences influence the internal belief-systems of those affected (Kube, 2023). In this context, future research could use experimental tasks from social psychology (e.g., Wirth, 2016) or game theory (e.g., Wang, Yang, Li, & Zhou, 2015), because these tasks often involve making decisions under uncertainty and can be easily formalized in terms of the aforementioned model (Eckert, Pawlowski, Rief, Endres, & Kirchner, 2023; Moutoussis et al., 2014; Smith et al., 2022). Moreover, these paradigms have demonstrated sensitivity to differences between individuals with depression and those without (Wang et al., 2015).

Understanding how the interplay of different biases in patients' internal belief-system leads to specific interpersonal experiences and behaviors, and how social observations influence these biases, has important implications for depression diagnostics and treatment. This is due to the likelihood that differences in generative models between patients (or patient subgroups) would necessitate corresponding differences in treatment. For example, patients harboring negative expectations about the hidden characteristics of an impending social situation (represented by the probability distribution $p(s_1)$) and low expectation of being able to change these characteristics through their actions (represented by the probability distribution $p\ (s_T|s_{T-1},\ \pi)$) may benefit most from acquiring social skills via social skills training (Bellack et al., 1996). Conversely, patients with similarly negative expectations but who anticipate being able to change these characteristics through hostile behavior (again represented by the probability distribution $p\ (s_T|s_{T-1},\ \pi)$) may first need interventions such as disciplined personal involvement (McCullough, 2006), interpersonal discrimination exercise (Mccullough, 2006), and behavioral experiments (Bennett-Levy et al., 2004) before they will be able to change their internal belief-systems and behaviors. Both examples illustrate that deeper understanding of the differences in patients' internal belief-systems can enhance depression diagnostics and facilitate its effective treatment. Therefore, future research could investigate whether estimating patients' internal belief systems, for example, by applying the above model to behavioral data from the proposed experimental tasks, could guide subsequent treatment.

Moreover, our approach offers new opportunities regarding the issue of persistent and treatment-resistant depression (Schramm, Klein, Elsaesser, Furukawa, & Domschke, 2020). It seems likely that therapies for these patients often fail to address key aspects of their internal belief-systems. For instance, no matter how many positive social experiences a patients encounters during behavioral activation, if they no longer believe that positive social observations are connected to the social environment's actual characteristics (as represented by the observational model $p(o_T|s_T)$), these experiences will hardly impact their internal belief-systems. Referring back to the example above, if a patient fails to infer any positive interpersonal states from observing smiles due to past negative interpersonal experiences, then being frequently smiled at will have little impact on their belief-system. Furthermore, these patients often expect to be unable to change the characteristics of their social environments through any action (referring to the transition probability $p(s_{T+1}|s_{T,}\ \pi)$). Therefore, future research should focus on how to effectively treat the observational model and self-efficacy expectations of these patients.

The heterogeneity of depression (Fried et al., 2022) and its interpersonal manifestations (Kupferberg & Hasler, 2023) make it challenging to explain the above-mentioned model in neurobiological terms. Depression, like many mental disorders (Eaton et al., 2023; Rief et al., 2023), cannot be "essentialized" (Adriaens & De Block, 2013) by a single, unified neural or biological pattern because of its complex, multicausal, and multifaceted nature (Gray, Müller, Eickhoff, & Fox, 2020; Marquand, Wolfers, Mennes, Buitelaar, & Beckmann, 2016; Pelin et al., 2021; Winter et al., 2023). Different hierarchical levels can affect interpersonal behaviors (Story et al., 2023), further complicating the issue. One way to better understand the neurobiological distortions of depressed individuals in social contexts would be to examine how systematic and finely-grained differences in patients' generative models correspond to neurobiological changes during a given interpersonal task. This might facilitate identifying less heterogeneous neural signatures within the spectrum of depressive disorders.

Our approach of characterizing different interpersonal phenomena within depression aligns with a broader trend of shifting away from fixed clinical categories and towards a more dimensional and transdiagnostic perspective on mental disorders (Conway, Forbes, & South, 2022; Dalgleish, Black, Johnston, & Bevan, 2020; Insel et al., 2010; Jungilligens, Paredes-Echeverri, Popkirov, Barrett, & Perez, 2022). Taking this perspective, distortions in the probability distribution $p(s_1)$ could signify not only low interpersonal trust and dysfunctional expectations in depression, but also across various mental disorders (Poggi, Richetin, & Preti, 2019; Rief et al., 2015) reflecting the impact of negative interpersonal experiences (Mauritz, Goossens, Draijer, & van Achterberg, 2013). Similarly, the observational model $p(o_T|s_T)$ might represent some perceptual and cognitive distortions observed across different mental disorders (Horga & Abi-Dargham, 2019; LeMoult & Gotlib, 2019; Van den Bergh, Brosschot, Critchley, Thayer, & Ottaviani, 2021). Lastly, biased transition probabilities $p\ (s_T|s_{T-1},\pi)$ could be relevant across mental disorders associated with learned helplessness or low social self-efficacy (Maier & Seligman, 2016; Niu et al., 2023).

Please note that our approach does not account for all aspects of depression, such as its episodic nature or related somatic symptoms (cf., Smith et al., 2018). This limitation also applies to the realm of interpersonal phenomena. For instance, developing models that describe the interactions between multiple agents with varying generative models could be beneficial. This would enable researchers to study dynamic interaction patterns between patients and healthy individuals, and the resulting impacts on the involved individuals' generative models.

Testable hypotheses for future research derived from our approach and a general outline of their empirical investigation are depicted in Table 1.

4. Conclusion

There is solid evidence that depression is associated with different distortions in interpersonal experience and behavior. We suggest that this heterogeneity can be attributed to the interplay of different biases in patients' internal belief-systems about their social world, leading to unfortunate conclusions about this world's genuine properties and corresponding distorted behaviors.

By formalizing these biases within the theory of active inference as a partially observable Markov decision process model, we enable researchers to describe interpersonal phenomena in depression more precisely, and to investigate their dynamic change in response to the social environment more systematically.

Future investigations can take our approach to identify systematic differences among patients with depression (or subgroups thereof) and to investigate how new social experiences influence the internal belief-systems of those affected. Future research should explore how our approach could augment depression diagnostics and personalize its treatment.

Use of generative AI in scientific writing

We declare that we used generative AI tools solely for text editing purposes (such as translation and spelling correction) in the preparation of this article. We did not use these tools to create any scientific content or search for scientific sources. All changes made were carefully reviewed for accuracy.

Funding

This work was funded by the Hessian Ministry of Higher Education, Research, Science and the Arts (Germany) as part of the cluster initiative "The Adaptive Mind".

Related links

A previous version of this manuscript is available on the PsyArXiv preprint server: https://psyarxiv.com/bp9re/. Neither the preprint nor the supplemental material were changed during the revision process.

CRediT authorship contribution statement

Lukas Kirchner: Conceptualization, Methodology, Project administration, Visualization, Writing – original draft. **Anna-Lena Eckert:** Conceptualization, Methodology, Supervision, Writing – review & editing. **Max Berg:** Supervision, Writing – review & editing. **Dominik Endres:** Supervision, Writing – review & editing, Methodology. **Benjamin Straube:** Supervision, Writing – review & editing. **Winfried Rief:** Funding acquisition, Supervision, Writing – review & editing.

Declaration of competing interest

This work was conducted as part of the cluster initiative "The Adaptive Mind" (TAM), which aims to bring together scientists from experimental psychology, clinical psychology, cognitive science, and artificial intelligence to enhance the understanding of how the human mind effectively adapts to changing conditions. This initiative also focuses on investigating the role of social inferences in the development of mental disorders.

Table 1
Testable hypotheses for future research derived from Fig. 1 and a general outline of their empirical investigation.

Distortions in external behaviors	Distortions in the internal belief-system	Expected biases in the generative model	Empirical investigation (closely based on Schwartenbeck et al., 2016)
Dependent behaviors	• Vague expectations regarding upcoming social states • Low confidence in social observations • High efficacy-expectations for dependent action sequences	• $p\ (s_1)$: Low probabilities for both negative and positive states • $p\ (o_1\|s_1)$: Low probabilities to infer negative or positive states from observations • $p\ (s_T\|s_{T-1}, \pi)$: High probabilities to transform negative states through dependent action sequences	1) Consider experimental tasks requiring multiple runs of interpersonal behavior (e.g., choices) under uncertainty based on social observations, and add both an information-seeking (e.g., exploration of the social environment) and a reward-seeking (e.g., monetary incentives) component.
Hostile behaviors	• Negative expectations regarding upcoming social states • Negative interpretation of social observations • High efficacy-expectations for hostile action sequences	• $p\ (s_1)$: High probabilities for negative states and low probabilities for positive states • $p\ (o_1\|s_1)$: High probabilites to infer negative states from observations • $p\ (s_T\|s_{T-1}, \pi)$: High probabilities to transform negative states through hostile action sequences	2) Ensure that these tasks contain action sequences that reflect the behavior of interest (e.g., hostile behavior) and associate these action sequences with a cost (e.g., loss of financial gain) 3) Specify a subjective generative model (e.g., a partially observable Markov decision process model) that predicts participants' responses during the task
Passive behaviors	• Vague expectations regarding upcoming social states • Low confidence in social observations • Low efficacy-expectations for all action sequences	• $p\ (s_1)$: Low probabilities for both negative and positive states • $p\ (o_1\|s_1)$: Low probabilities to infer negative or positive states from social observations • $p\ (s_T\|s_{T-1}, \pi)$: Low probability to transform negative social states through action sequences	4) Simulate task behavior to investigate whether the tasks produce relevant differences in behavior given the generative model 5) Invert the generative model based on participants' real task behavior to recover its parameters
Avoidant behaviors	• Negative expectations regarding upcoming social states • Negative interpretation of social observations • Low efficacy-expectations for all action sequences	• $p\ (s_1)$: High probabilities for negative states and low probabilities for positive states • $p\ (o_1\|s_1)$: High probabilities to infer negative states from observations • $p\ (s_T\|s_{T-1}, \pi)$: Low probabilities to transform negative social states through action sequences	6) Investigate between-subject or group differences and compare different plausible models of the data 7) Evaluate your model of choice by fitting its parameters using a training sample and testing its accuracy in a test sample

Data availability

No data was used for the research described in the article.

Acknowledgments

I would like to thank all the collaborators of the cluster initiative "The Adaptive Mind" (TAM) for the productive discussions that made this project possible.

References

Adams, R. A., Shipp, S., & Friston, K. J. (2013). Predictions not commands: Active inference in the motor system. *Brain Structure and Function, 218*(3), 611–643. https://doi.org/10.1007/s00429-012-0475-5

Adriaens, P. R., & De Block, A. (2013). Why we essentialize mental disorders. *Journal of Medicine and Philosophy*. https://doi.org/10.1093/jmp/jht008

Attias, H. (2003). Planning by probabilistic inference. In C. M. Bishop, & B. J. Frey (Eds.), *Proceedings of the ninth international workshop on artificial intelligence and statistics, R4* pp. 9–16). PMLR. https://proceedings.mlr.press/r4/attias03a.html.

Badcock, P. B., Davey, C. G., Whittle, S., Allen, N. B., & Friston, K. J. (2017). The depressed brain: An evolutionary systems theory. *Trends in Cognitive Sciences, 21*(3), 182–194. https://doi.org/10.1016/j.tics.2017.01.005

Baldwin, M. W. (1992). Relational schemas and the processing of social information. *Psychological Bulletin, 112*(3), 461–484. https://doi.org/10.1037/0033-2909.112.3.461

Barkus, E., & Badcock, J. C. (2019). A transdiagnostic perspective on social anhedonia. *Frontiers in Psychiatry, 10*, 216. https://doi.org/10.3389/fpsyt.2019.00216

Beck, A. T. (1963). Thinking and depression: I. Idiosyncratic content and cognitive distortions. *Archives of General Psychiatry, 9*(4), 324–333. https://doi.org/10.1001/ARCHPSYC.1963.01720160014002

Beck, A. T. (1964). Thinking and depression: II. Theory and therapy. *Archives of General Psychiatry, 10*(6), 561. https://doi.org/10.1001/archpsyc.1964.01720240015003

Beeson, C. M. L., Brittain, H., & Vaillancourt, T. (2020). The temporal precedence of peer rejection, rejection sensitivity, depression, and aggression across adolescence. *Child Psychiatry and Human Development, 51*(5), 781–791. https://doi.org/10.1007/s10578-020-01008-2

Bellack, A. S., Hersen, M., & Himmelhoch, J. M. (1996). Social skills training for depression. In *Sourcebook of psychological treatment manuals for adult disorders* (pp. 179–200). Springer US. https://doi.org/10.1007/978-1-4899-1528-3_5.

Bennett-Levy, J., Butler, G., Fennell, M., Hackmann, A., Mueller, M., & Westbrook, D. (2004). *Oxford guide to behavioural experiments in cognitive therapy*. Oxford University Press. https://doi.org/10.1093/med:psych/9780198529163.001.0001

Berg, M., Feldmann, M., Kirchner, L., & Kube, T. (2022). Oversampled and undersolved: Depressive rumination from an active inference perspective. *Neuroscience & Biobehavioral Reviews, 142*(August), Article 104873. https://doi.org/10.1016/j.neubiorev.2022.104873

Berger, J. O. (2011). *Statistical decision theory and Bayesian analysis*. Springer.

Bird, T., Tarsia, M., & Schwannauer, M. (2018). Interpersonal styles in major and chronic depression: A systematic review and meta-analysis. *Journal of Affective Disorders, 239*, 93–101. https://doi.org/10.1016/j.jad.2018.05.057

Botvinick, M., & Toussaint, M. (2012). Planning as inference. *Trends in Cognitive Sciences, 16*(10), 485–488. https://doi.org/10.1016/j.tics.2012.08.006

Bourke, C., Douglas, K., & Porter, R. (2010). Processing of facial emotion expression in major depression: A review. *Australian and New Zealand Journal of Psychiatry, 44*(8), 681–696. https://doi.org/10.3109/00048674.2010.496359

Brown, L. D. (1981). A complete class theorem for statistical problems with finite sample spaces. *Annals of Statistics, 9*(6), 1289–1300. https://doi.org/10.1214/aos/1176345645

Cain, N. M., Ansell, E. B., Wright, A. G. C., Hopwood, C. J., Thomas, K. M., Pinto, A., et al. (2012). Interpersonal pathoplasticity in the course of major depression. *Journal of Consulting and Clinical Psychology, 80*(1), 78–86. https://doi.org/10.1037/a0026433

Chekroud, A. M. (2015). Unifying treatments for depression: An application of the free energy principle. *Frontiers in Psychology, 6*(FEB), 153. https://doi.org/10.3389/FPSYG.2015.00153/BIBTEX

Chen, C., Takahashi, T., Nakagawa, S., Inoue, T., & Kusumi, I. (2015). Reinforcement learning in depression: A review of computational research. *Neuroscience & Biobehavioral Reviews, 55*, 247–267. https://doi.org/10.1016/j.neubiorev.2015.05.005

Conway, C. C., Forbes, M. K., & South, S. C. (2022). A Hierarchical Taxonomy of Psychopathology (HiTOP) primer for mental health researchers. *Clinical Psychological Science, 10*(2), 236–258. https://doi.org/10.1177/21677026211017834

Coyne, J. C. (1976). Toward an interactional description of depression. *Psychiatry, 39*(1), 28–40. https://doi.org/10.1080/00332747.1976.11023874

Cuijpers, P., Muñoz, R. F., Clarke, G. N., & Lewinsohn, P. M. (2009). Psychoeducational treatment and prevention of depression: The "coping with depression" course thirty years later. *Clinical Psychology Review, 29*(5), 449–458. https://doi.org/10.1016/j.cpr.2009.04.005

Da Costa, L., Parr, T., Sajid, N., Veselic, S., Neacsu, V., & Friston, K. (2020). Active inference on discrete state-spaces: A synthesis. *Journal of Mathematical Psychology, 99*, Article 102447. https://doi.org/10.1016/j.jmp.2020.102447

Da Costa, L., Parr, T., Sengupta, B., & Friston, K. (2021). Neural dynamics under active inference: Plausibility and efficiency of information processing. *Entropy, 23*(4), 454. https://doi.org/10.3390/e23040454

Dalgleish, T., Black, M., Johnston, D., & Bevan, A. (2020). Transdiagnostic approaches to mental health problems: Current status and future directions. *Journal of Consulting and Clinical Psychology, 88*(3), 179–195. https://doi.org/10.1037/ccp0000482

Dawood, S., Thomas, K. M., Wright, A. G. C., & Hopwood, C. J. (2013). Heterogeneity of interpersonal problems among depressed young adults: Associations with substance abuse and pathological personality traits. *Journal of Personality Assessment, 95*(5), 513–522. https://doi.org/10.1080/00223891.2013.781031

de Jonge, P., Wardenaar, K. J., & Wichers, M. (2015). What kind of thing is depression? *Epidemiology and Psychiatric Sciences, 24*(4), 312–314. https://doi.org/10.1017/S2045796015000207

Dobson, K. S., Quigley, L., & Dozois, D. J. A. (2014). Toward an integration of interpersonal risk models of depression and cognitive-behaviour therapy. *Australian Psychologist, 49*(6), 328–336. https://doi.org/10.1111/ap.12079

Eaton, N. R., Bringmann, L. F., Elmer, T., Fried, E. I., Forbes, M. K., Greene, A. L., et al. (2023). A review of approaches and models in psychopathology conceptualization research. *Nature Reviews Psychology, 2*(10), 622–636. https://doi.org/10.1038/s44159-023-00218-4

Eckert, A.-L., Pawlowski, J., Rief, W., Endres, D., & Kirchner, L. (2023). Simulating active inference of interpersonal context within and across mental disorders. https://doi.org/10.31234/osf.io/8aexf.

Everaert, J. (2021). Interpretation of ambiguity in depression. *Current Opinion in Psychology, 41*, 9–14. https://doi.org/10.1016/j.copsyc.2021.01.003

Everaert, J., Podina, I. R., & Koster, E. H. W. (2017). A comprehensive meta-analysis of interpretation biases in depression. *Clinical Psychology Review, 58*, 33–48. https://doi.org/10.1016/j.cpr.2017.09.005

Evraire, L. E., & Dozois, D. J. A. (2011). An integrative model of excessive reassurance seeking and negative feedback seeking in the development and maintenance of depression. *Clinical Psychology Review, 31*(8), 1291–1303. https://doi.org/10.1016/j.cpr.2011.07.014

Fried, E. I., Flake, J. K., & Robinaugh, D. J. (2022). Revisiting the theoretical and methodological foundations of depression measurement. *Nature Reviews Psychology, 1*(6), 358–368. https://doi.org/10.1038/s44159-022-00050-2

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience, 11*(2), 127–138. https://doi.org/10.1038/nrn2787

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., O'Doherty, J., & Pezzulo, G. (2016). Active inference and learning. *Neuroscience & Biobehavioral Reviews, 68*, 862–879. https://pubmed.ncbi.nlm.nih.gov/27375276/.

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2017). Active inference: A process theory. *Neural Computation, 29*(1), 1–49. https://doi.org/10.1162/NECO_a_00912

Gadassi, R., & Rafaeli, E. (2015). Interpersonal perception as a mediator of the depression–interpersonal difficulties link: A review. *Personality and Individual Differences, 87*, 1–7. https://doi.org/10.1016/j.paid.2015.07.023

Gamble, B., Moreau, D., Tippett, L. J., & Addis, D. R. (2019). Specificity of future thinking in depression: A meta-analysis. *Perspectives on Psychological Science, 14*(5), 816–834. https://doi.org/10.1177/1745691619851784

Gao, S., Assink, M., Cipriani, A., & Lin, K. (2017). Associations between rejection sensitivity and mental health outcomes: A meta-analytic review. *Clinical Psychology Review, 57*, 59–74. https://doi.org/10.1016/j.cpr.2017.08.007

Gariépy, G., Honkaniemi, H., & Quesnel-Vallée, A. (2016). Social support and protection from depression: Systematic review of current findings in western countries. *British Journal of Psychiatry, 209*(4), 284–293. https://doi.org/10.1192/bjp.bp.115.169094

Goldberg, D. (2011). The heterogeneity of "major depression.". *World Psychiatry, 10*(3), 226–228. https://doi.org/10.1002/j.2051-5545.2011.tb00061.x

Gopnik, A. (2020). Childhood as a solution to explore–exploit tensions. *Philosophical Transactions of the Royal Society B: Biological Sciences, 375*(1803), Article 20190502. https://doi.org/10.1098/rstb.2019.0502

Gray, J. P., Müller, V. I., Eickhoff, S. B., & Fox, P. T. (2020). Multimodal abnormalities of brain structure and function in major depressive disorder: A meta-analysis of neuroimaging studies. *American Journal of Psychiatry, 177*(5), 422–434. https://doi.org/10.1176/appi.ajp.2019.19050560

Halahakoon, D. C., Kieslich, K., O'Driscoll, C., Nair, A., Lewis, G., & Roiser, J. P. (2020). Reward-processing behavior in depressed participants relative to healthy volunteers: A systematic reviews and meta-analysis. *JAMA Psychiatry, 77*(12), 1286. https://doi.org/10.1001/jamapsychiatry.2020.2139

Hames, J. L., Hagan, C. R., & Joiner, T. E. (2013). Interpersonal processes in depression. *Annual Review of Clinical Psychology, 9*, 355–377. https://doi.org/10.1146/annurev-clinpsy-050212-185553

Hauser, T. U., Skvortsova, V., De Choudhury, M., & Koutsouleris, N. (2022). The promise of a model-based psychiatry: Building computational models of mental ill health. *The Lancet Digital Health, 4*(11), e816–e828. https://doi.org/10.1016/S2589-7500(22)00152-2

Heins, C., Millidge, B., Demekas, D., Klein, B., Friston, K., Couzin, I. D., et al. (2022). pymdp: A Python library for active inference in discrete state spaces. *Journal of Open Source Software, 7*(73), 4098. https://doi.org/10.21105/joss.04098

Hirschfeld, R. M. A., Montgomery, S. A., Keller, M. B., Kasper, S., Schatzberg, A. F., Moller, H.-J., et al. (2000). Social functioning in depression. *The Journal of Clinical Psychiatry, 61*(4), 268–275. https://doi.org/10.4088/JCP.v61n0405

Horga, G., & Abi-Dargham, A. (2019). An integrative framework for perceptual disturbances in psychosis. *Nature Reviews Neuroscience, 20*(12), 763–778. https://doi.org/10.1038/s41583-019-0234-1

Huys, Q. J. M., Guitart-Masip, M., Dolan, R. J., & Dayan, P. (2015). Decision-theoretic psychiatry. *Clinical Psychological Science, 3*(3), 400–421. https://doi.org/10.1177/2167702614562040
Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., et al. (2010). Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. *American Journal of Psychiatry, 167*(7), 748–751. https://doi.org/10.1176/appi.ajp.2010.09091379
Jungilligens, J., Paredes-Echeverri, S., Popkirov, S., Barrett, L. F., & Perez, D. L. (2022). A new science of emotion: Implications for functional neurological disorder. *Brain: Journal of Neurology, 145*(8), 2648–2663. https://doi.org/10.1093/brain/awac204
Kaplan, R., & Friston, K. J. (2018). Planning and navigation as active inference. *Biological Cybernetics, 112*(4), 323–343. https://doi.org/10.1007/s00422-018-0753-2
Kavanagh, D. J. (2014). Self-efficacy and depression. In R. Schwarzer (Ed.), *Self-efficacy: Thought control of action* (2014th ed., pp. 177–194). Routledge.
Kirchner, L., Schummer, S. E., Krug, H., Kube, T., & Rief, W. (2022). How social rejection expectations and depressive symptoms bi-directionally predict each other – a cross-lagged panel analysis. *Psychology and Psychotherapy: Theory, Research and Practice, 95*(2), 477–492. https://doi.org/10.1111/papt.12383
Klerman, G. L., & Weissman, M. M. (1994). Interpersonal psychotherapy of depression: A brief, focused, specific strategy. https://rowman.com/ISBN/9781568213507/Interpersonal-Psychotherapy-of-Depression-A-Brief-Focused-Specific-Strategy.
Kouvaris, K., Clune, J., Kounios, L., Brede, M., & Watson, R. A. (2015). How evolution learns to generalise: Principles of under-fitting, over-fitting and induction in the evolution of developmental organisation. *Nature, 388*, 539–547. https://doi.org/10.48550/arXiv.1508.06854
Kube, T. (2023). Biased belief updating in depression. *Clinical Psychology Review, 103* (May), Article 102298. https://doi.org/10.1016/j.cpr.2023.102298
Kube, T., Schwarting, R., Rozenkrantz, L., Glombiewski, J. A., & Rief, W. (2020). Distorted cognitive processes in major depression: A predictive processing perspective. *Biological Psychiatry, 87*(5), 388–398. https://doi.org/10.1016/j.biopsych.2019.07.017
Kupferberg, A., Bicks, L., & Hasler, G. (2016). Social functioning in major depressive disorder. *Neuroscience & Biobehavioral Reviews, 69*, 313–332. https://doi.org/10.1016/j.neubiorev.2016.07.002
Kupferberg, A., & Hasler, G. (2023). The social cost of depression: Investigating the impact of impaired social emotion regulation, social cognition, and interpersonal behavior on social functioning. *Journal of Affective Disorders Reports, 14*(December 2022), Article 100631. https://doi.org/10.1016/j.jadr.2023.100631
LeMoult, J., & Gotlib, I. H. (2019). Depression: A cognitive perspective. *Clinical Psychology Review, 69*, 51–66. https://doi.org/10.1016/j.cpr.2018.06.008
Lindley, D. V. (1956). On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics, 27*(4), 986–1005. https://doi.org/10.1214/aoms/1177728069
Lipsitz, J. D., & Markowitz, J. C. (2013). Mechanisms of change in interpersonal therapy (IPT). *Clinical Psychology Review, 33*(8), 1134–1147. https://doi.org/10.1016/j.cpr.2013.09.002
Liu, R. T., & Alloy, L. B. (2010). Stress generation in depression: A systematic review of the empirical literature and recommendations for future study. *Clinical Psychology Review, 30*(5), 582–593. https://doi.org/10.1016/j.cpr.2010.04.010
MacKay, D. J. C. (1992). Information-based objective functions for active data selection. *Neural Computation, 4*(4), 590–604. https://doi.org/10.1162/neco.1992.4.4.590
Maier, S. F., & Seligman, M. E. P. (2016). Learned helplessness at fifty: Insights from neuroscience. *Psychological Review, 123*(4), 349–367. https://doi.org/10.1037/rev0000033
Marquand, A. F., Wolfers, T., Mennes, M., Buitelaar, J., & Beckmann, C. F. (2016). Beyond lumping and splitting: A review of computational approaches for stratifying psychiatric disorders. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 1*(5), 433–447. https://doi.org/10.1016/j.bpsc.2016.04.002
Mauritz, M. W., Goossens, P. J. J., Draijer, N., & van Achterberg, T. (2013). Prevalence of interpersonal trauma exposure and trauma-related disorders in severe mental illness. *European Journal of Psychotraumatology, 4*(1), 1–15. https://doi.org/10.3402/ejpt.v4i0.19985
McCullough, J. P. (2003). Treatment for chronic depression using cognitive behavioral analysis system of psychotherapy (CBASP). *Journal of Clinical Psychology, 59*(8), 833–846. https://doi.org/10.1002/jclp.10176
Mccullough, J. P. (2006). *Treating chronic depression with disciplined personal involvement: Cognitive behavioral analysis system of psychotherapy (CBASP)*. Springer.
McCullough, J. P. (2006). Healing interpersonal trauma using the interpersonal discrimination exercise. In *Treating chronic depression with disciplined personal involvement* (pp. 123–159). Springer US. https://doi.org/10.1007/978-0-387-31066-4_6.
McCullough, J., Schramm, E., & Penberthy, J. K. (2014). CBASP as a distinctive treatment for persistent depressive disorder. In *CBASP as a distinctive treatment for persistent depressive disorder*. Routledge. https://doi.org/10.4324/9781315743196.
Monroe, S. M., & Anderson, S. F. (2015). Depression: The shroud of heterogeneity. *Current Directions in Psychological Science, 24*(3), 227–231. https://doi.org/10.1177/0963721414568342
Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. *Trends in Cognitive Sciences, 16*(1), 72–80. https://doi.org/10.1016/j.tics.2011.11.018
Moutoussis, M., Trujillo-Barreto, N. J., El-Deredy, W., Dolan, R. J., & Friston, K. J. (2014). A formal model of interpersonal inference. *Frontiers in Human Neuroscience, 8* (MAR), 160. https://doi.org/10.3389/fnhum.2014.00160
Muñoz, R. F., Ghosh Ippen, C., Rao, S., Le, H.-N., & Valdes Dwyer, E. (2000). *Manual for group cognitive-behavioral therapy of major depression*.
Nagy, E., & Moore, S. (2017). Social interventions: An effective approach to reduce adult depression? *Journal of Affective Disorders, 218*, 131–152. https://doi.org/10.1016/j.jad.2017.04.043
Nanni, V., Uher, R., & Danese, A. (2012). Childhood maltreatment predicts unfavorable course of illness and treatment outcome in depression: A meta-analysis. *American Journal of Psychiatry, 169*(2), 141–151. https://doi.org/10.1176/appi.ajp.2011.11020335
Nestor, B. A., Sutherland, S., & Garber, J. (2022). Theory of mind performance in depression: A meta-analysis. *Journal of Affective Disorders, 303*, 233–244. https://doi.org/10.1016/j.jad.2022.02.028
Ng, T. H., Alloy, L. B., & Smith, D. V. (2019). Meta-analysis of reward processing in major depressive disorder reveals distinct abnormalities within the reward circuit. *Translational Psychiatry, 9*(1), 293. https://doi.org/10.1038/s41398-019-0644-x
Niu, G., Shi, X., Yao, L., Yang, W., Jin, S., & Xu, L. (2023). Social exclusion and depression among undergraduate students: The mediating roles of rejection sensitivity and social self-efficacy. *Current Psychology, 42*(28), 24198–24207. https://doi.org/10.1007/s12144-022-03318-1
Pelin, H., Ising, M., Stein, F., Meinert, S., Meller, T., Brosch, K., et al. (2021). Identification of transdiagnostic psychiatric disorder subtypes using unsupervised learning. *Neuropsychopharmacology, 46*(11), 1895–1905. https://doi.org/10.1038/s41386-021-01051-0
Platt, B., Kadosh, K. C., & Lau, J. Y. F. (2013). The role of peer rejection in adolescent depression. *Depression and Anxiety, 30*(9), 809–821. https://doi.org/10.1002/da.22120
Poggi, A., Richetin, J., & Preti, E. (2019). Trust and rejection sensitivity in personality disorders. *Current Psychiatry Reports, 21*(8), 69. https://doi.org/10.1007/s11920-019-1059-3
Porcelli, S., Van Der Wee, N., van der Werff, S., Aghajani, M., Glennon, J. C., van Heukelum, S., et al. (2019). Social brain, social dysfunction and social withdrawal. *Neuroscience & Biobehavioral Reviews, 97*, 10–33. https://doi.org/10.1016/j.neubiorev.2018.09.012
Renner, F., Arntz, A., Leeuw, I., & Huibers, M. (2013). Treatment for chronic depression using schema therapy. *Clinical Psychology: Science and Practice, 20*(2), 166–180. https://doi.org/10.1111/cpsp.12032
Rief, W., Glombiewski, J. A., Gollwitzer, M., Schubö, A., Schwarting, R., & Thorwart, A. (2015). Expectancies as core features of mental disorders. *Current Opinion in Psychiatry, 28*(5), 378–385. https://doi.org/10.1097/YCO.0000000000000184
Rief, W., Hofmann, S. G., Berg, M., Forbes, M. K., Pizzagalli, D. A., Zimmermann, J., et al. (2023). Do we need a novel framework for classifying psychopathology? - a discussion paper. *Clinical Psychology in Europe, 5*(4). https://doi.org/10.32872/cpe.11699
Rief, W., & Joormann, J. (2019). Revisiting the cognitive model of depression: The role of expectations. *Clinical Psychology in Europe, 1*(1). https://doi.org/10.32872/cpe.v1i1.32605
Roediger, E., Stevens, B. A., & Brockman, R. (2018). *Contextual schema therapy: An integrative approach to personality disorders, emotional dysregulation, and interpersonal functioning*. New Harbinger Publications. https://psycnet.apa.org/record/2018-30793-000.
Roepke, A. M., & Seligman, M. E. P. (2016). Depression and prospection. *British Journal of Clinical Psychology, 55*(1), 23–48. https://doi.org/10.1111/bjc.12087
Safra, L., Chevallier, C., & Palminteri, S. (2019). Depressive symptoms are associated with blunted reward learning in social contexts. *PLoS Computational Biology, 15*(7), Article e1007224. https://doi.org/10.1371/journal.pcbi.1007224
Schramm, E., Klein, D. N., Elsaesser, M., Furukawa, T. A., & Domschke, K. (2020). Review of dysthymia and persistent depressive disorder: History, correlates, and clinical implications. *The Lancet Psychiatry, 7*(9), 801–812. https://doi.org/10.1016/S2215-0366(20)30099-7
Schwartenbeck, P., FitzGerald, T. H. B., Mathys, C., Dolan, R., Wurst, F., Kronbichler, M., et al. (2015). Optimal inference with suboptimal models: Addiction and active bayesian inference. *Medical Hypotheses, 84*(2), 109–117. https://doi.org/10.1016/j.mehy.2014.12.007
Schwartenbeck, P., & Friston, K. (2016). Computational phenotyping in psychiatry: A worked example. *Eneuro, 3*(4). https://doi.org/10.1523/ENEURO.0049-16.2016. ENEURO.0049-16.2016.
Schwartenbeck, P., Passecker, J., Hauser, T. U., FitzGerald, T. H., Kronbichler, M., & Friston, K. J. (2019). Computational mechanisms of curiosity and goal-directed exploration. *Elife, 8*, Article e41703. https://doi.org/10.7554/eLife.41703
Segrin, C. (1990). A meta-analytic review of social skill deficits in depression. *Communication Monographs, 57*(4), 292–308. https://doi.org/10.1080/03637759009376204
Segrin, C. (2000). Social skills deficits associated with depression. *Clinical Psychology Review, 20*(3), 379–403. https://doi.org/10.1016/S0272-7358(98)00104-4
Seligman, M. E. P. (1972). Learned helplessness. *Annual Review of Medicine, 23*(1), 407–412. https://doi.org/10.1146/annurev.me.23.020172.002203
Simon, S., Cain, N. M., Wallner Samstag, L., Meehan, K. B., & Muran, J. C. (2015). Assessing interpersonal subtypes in depression. *Journal of Personality Assessment, 97* (4), 364–373. https://doi.org/10.1080/00223891.2015.1011330
Smith, R., Alkozei, A., Killgore, W. D. S., & Lane, R. D. (2018). Nested positive feedback loops in the maintenance of major depression: An integration and extension of previous models. *Brain, Behavior, and Immunity, 67*, 374–397. https://pubmed.ncbi.nlm.nih.gov/28943294/.
Smith, R., Badcock, P., & Friston, K. J. (2021). Recent advances in the application of predictive coding and active inference models within clinical neuroscience. *Psychiatry and Clinical Neurosciences, 75*(1), 3–13. https://doi.org/10.1111/pcn.13138

Smith, R., Friston, K. J., & Whyte, C. J. (2022). A step-by-step tutorial on active inference and its application to empirical data. *Journal of Mathematical Psychology, 107*, Article 102632. https://doi.org/10.1016/j.jmp.2021.102632

Smith, R., Kirlic, N., Stewart, J. L., Touthang, J., Kuplicki, R., Khalsa, S. S., et al. (2021). Greater decision uncertainty characterizes a transdiagnostic patient sample during approach-avoidance conflict: A computational modelling approach. *Journal of Psychiatry & Neuroscience, 46*(1), E74–E87. https://doi.org/10.1503/jpn.200032

Smith, R., Schwartenbeck, P., Stewart, J. L., Kuplicki, R., Ekhtiari, H., & Paulus, M. P. (2020). Imprecise action selection in substance use disorder: Evidence for active learning impairments when solving the explore-exploit dilemma. *Drug and Alcohol Dependence, 215*(April). https://doi.org/10.1016/j.drugalcdep.2020.108208

Starr, L. R., & Davila, J. (2008). Excessive reassurance seeking, depression, and interpersonal rejection: A meta-analytic review. *Journal of Abnormal Psychology, 117*(4), 762–775. https://doi.org/10.1037/a0013866

Story, G. W., Smith, R., Moutoussis, M., Berwian, I. M., Nolte, T., Bilek, E., et al. (2023). A social inference model of idealization and devaluation. *Psychological Review*. https://doi.org/10.1037/rev0000430

Suffel, A., Nagels, A., Steines, M., Kircher, T., & Straube, B. (2020). Feeling addressed! - the neural processing of social communicative cues in patients with major depression. *Human Brain Mapping, 41*(13), 3541–3554. https://doi.org/10.1002/hbm.25027

Teichman, Y., & Teichman, M. (1990). Interpersonal view of depression: Review and integration. *Journal of Family Psychology, 3*(4), 349–367. https://doi.org/10.1037/h0080549

Tse, W. S., & Bond, A. J. (2004). The impact of depression on social skills: A review. *The Journal of Nervous and Mental Disease, 192*(4), 260–268. https://doi.org/10.1097/01.nmd.0000120884.60002.2b

Uphoff, E., Ekers, D., Robertson, L., Dawson, S., Sanger, E., South, E., et al. (2020). Behavioural activation therapy for depression in adults. *Cochrane Database of Systematic Reviews, 2020*(7), CD013305. https://doi.org/10.1002/14651858.CD013305.PUB2/FULL

Van den Bergh, O., Brosschot, J., Critchley, H., Thayer, J. F., & Ottaviani, C. (2021). Better safe than sorry: A common signature of general vulnerability for psychopathology. *Perspectives on Psychological Science, 16*(2), 225–246. https://doi.org/10.1177/1745691620950690

Wald, A. (1947). An essentially complete class of admissible decision functions. *The Annals of Mathematical Statistics, 18*(4), 549–555. https://doi.org/10.1214/aoms/1177730345

Wang, Y., Yang, L. Q., Li, S., & Zhou, Y. (2015). Game theory paradigm: A new tool for investigating social dysfunction in major depressive disorders. *Frontiers in Psychiatry, 6*(SEP), 128. https://doi.org/10.3389/FPSYT.2015.00128/BIBTEX

Weightman, M. J., Air, T. M., & Baune, B. T. (2014). A review of the role of social cognition in major depressive disorder. *Frontiers in Psychiatry, 5*(NOV), 179. https://doi.org/10.3389/FPSYT.2014.00179/BIBTEX

Wichers, M. (2014). The dynamic nature of depression: A new micro-level perspective of mental disorder that meets current challenges. *Psychological Medicine, 44*(7), 1349–1360. https://doi.org/10.1017/S0033291713001979

Winter, N. R., Blanke, J., Leenings, R., Ernsting, J., Fisch, L., Sarink, K., et al. (2023). A systematic evaluation of machine learning-based biomarkers for major depressive disorder across modalities. *medRxiv*. https://doi.org/10.1101/2023.02.27.23286311, 2023.02.27.23286311.

Wirth, J. H. (2016). Methods for investigating social exclusion. In *Social exclusion* (pp. 25–47). Springer International Publishing. https://doi.org/10.1007/978-3-319-33033-4_2.

Yang, X., Su, Y., Yang, F., Song, Y., Yan, J., Luo, Y., et al. (2022). Neurofunctional mapping of reward anticipation and outcome for major depressive disorder: A voxel-based meta-analysis. *Psychological Medicine, 52*(15), 3309–3322. https://doi.org/10.1017/S0033291722002707