Deviations from normative emotion representation predicts social and affective dysfunction

Authors: Yi Yang Teoh[1], & Oriel FeldmanHall[1,2]*
[1]Department of Cognitive and Psychological Sciences, Brown University, Providence, RI 02912
[2]Robert J. and Nancy D. Carney Institute for Brain Science, Brown University, Providence, RI 02912

*Corresponding Author:
Oriel FeldmanHall
Brown University
190 Thayer St. Providence, RI 02912
oriel.feldmanhall@brown.edu

Author contributions: Conceptualization: Y.Y.T., and O.F.H. Formal Analysis: Y.Y.T Funding Acquisition: O.F.H. Investigation: Y.Y.T. Methodology: Y.Y.T. Software Y.Y.T. Supervision: O.F.H. Validation: Y.Y.T. Writing – original draft: Y.Y.T., and O.F.H. Writing – review and editing: Y.Y.T., and O.F.H.

Acknowledgements: We would like to thank Joseph Heffner for his assistance in data management. This work is supported by a Banting Postdoctoral Fellowship FRN 198553 from the Social Sciences and Humanities Research Council (SSHRC) awarded to Y.Y.T. and. a Center of Biological Research Excellence P20GM103645 from the National Institute of General Medical Sciences awarded to O.F.H.

Abstract

Emotion concepts serve as key guideposts for understanding our feelings, communicating them to others, and selecting appropriate actions. Yet people differ in how they represent these concepts—'anger' may mean rage for one person and mild irritation for another—even though we typically assume a shared understanding. Because emotion concepts support efficient communication and coordination, such divergences raise a critical question: what happens when an individual's emotion-concept representations deviate from the group norm? Here we investigate an overlooked aspect of emotion, the degree to which conceptual representations of emotion integrate and align with the normative emotional scaffolding that emerges at the population level. We find across three studies (N=1,490) that greater misalignment from the group-level representation of emotion concepts has broad costs for adaptive social and emotional functioning. Individuals whose concepts are more misaligned are less sensitive to others' antisocial behaviors during economic games, leading to less strategic punishment of norm violators and poorer social coordination. Consistent with the idea that emotion unfolds between people and the larger social community, greater misalignment also tracks pervasive affective dysfunction, including more severe alexithymia, anhedonia, and depression. That emotion misalignment may be a key through-line or pathway disrupting affective functioning seen in mood disorders, offers new theoretical insight into the possible mechanisms underlying emotion processing. Together, these findings underscore that emotion knowledge is fundamentally social and that systematic deviations from shared emotion concepts undermine both social adaptation and emotional health.

Main

Emotions are a central feature of human life[1] and vital for personal[2–5] and social well-being[6–8]. Because everyone reports feeling emotions like sadness, joy, and anger, it proffers the illusion that humans share a similar understanding of such emotion concepts[9]. Emotion concepts, such as 'anger', are socially constructed labels of subjective experiences[10–13]. We learn what 'anger' means based on social norms and by observing how others use the term and respond to it[14]. At the same time, concepts vary considerably across individuals; anger may resemble quiet frustration for some and hot rage for others. This paradox between shared norms and individual idiosyncrasies exposes a key gap in our understanding of emotion: What is gained by possessing a normative representation of emotion concepts? By characterizing individuals' representations of emotion concepts and quantifying their alignment with population-level representations, we test the consequences of having a representation of emotion knowledge that deviates from the norm.

The idea that emotion concepts are universally held traces back to early theories positing that a small set of basic emotions, selected by evolutionary pressures, coordinate social behavior[15,16]. The premise is that widely shared emotion concepts make others' mental states more legible and predictable, enabling the efficient communication and perspective-taking that is required to sustain cooperation. Contemporary work similarly argues that many emotion concepts are broadly shared across cultures[17–23]. Despite evidence of these common structural properties, emotion concepts also vary across individuals, revealing that emotion concepts have fuzzy boundaries[24–28]. These individual differences in emotion knowledge—known as, emotion granularity or differentiation—capture variations in people's capacity to make fine-grained distinctions among similar emotions, an ability linked to more adaptive responding[29] (i.e., poorer differentiation is associated with greater affective dysfunction[28–31] seen in alexithymia and depression[32–36]).

While emotion differentiation provides one explanatory mechanism for affective dysfunction, the sole focus on an individual's ability to differentiate between emotion concepts ignores the fundamentally social aspect of emotion knowledge. In everyday life, we learn what emotions are not just through introspection, but by observing how others label, express, and respond to emotions. Emotion differentiation measures, however, are blind to whether two people *share* a similar understanding of emotion concepts. For example, while two people might be very good at differentiating between emotion concepts, they may not conceptualize 'anger' in similar ways. The possibility of significant heterogeneity in emotion knowledge challenges the usual assumption that emotion concepts serve as efficient conduits for transmitting information[37]. Part of the power of emotion labels is that they compress widely shared cultural knowledge into compact signals[38,39], signals used to communicate with the world[40], predict the actions of others[41], and guide our own responses[12,27,42]. If, however, shared emotion knowledge is what enables social coordination[41], communication[43], and connection[6], then any deviations from those shared meanings should make those social functions far harder.

To test whether deviations from shared emotion-concept knowledge actually undermines social functioning, we measure the degree to which an individual's emotion knowledge is aligned with the group. To make it concrete, imagine telling your friends that you feel sad, a term which strips away much of the granularity of your actual internal experience. If your friends share a common conceptual understanding of 'sadness', it will help them infer your actual affective experience and thus lead to better communication and coordination. If instead, you possess a

highly misaligned representation where 'sadness' refers to a class of experiences more akin to what others might understand as 'anger', such misalignment would likely fuel misinterpretations of your actual affective state. Because emotion concepts are socially defined and used to scaffold adaptive social behaviors, we predict that greater misalignment will make it harder for individuals to draw on the normative schemas embedded in emotion knowledge—encumbering social coordination and hindering judgments like whom to trust and when to punish norm violators.

Alignment of one's emotion concepts with the group likely confers other benefits as well. Because emotions unfold within the context of friends, family, and the broader community[41,44,45], any misalignment likely disrupts how one affectively responds and interprets the interpersonal dynamics occurring with others. Disrupted social dynamics (two people misunderstanding one another), likely exacerbates negative affect, and the known attendant symptoms associated with heightened negative mood states (e.g., depression). Put simply, misalignment might have far broader consequences than only influencing uncoordinated social behaviors and may be one of the pathways driving mental health disorders marked by affective dysfunction.

To probe participants' internal representations of emotion concepts across three studies (N_{total} = 1,490), we use the Emotion Classification Task (Fig 1A) to explicitly map out emotion concepts along the core affective dimensions of valence and arousal[27,46]. Participants are asked to draw on their emotion knowledge to place 20 distinct emotion labels (e.g. angry, happy, sad, etc.) within a 500- by 500-pixelated grid defined by valence (x-axis) and arousal (y-axis). To characterize the alignment of participants' representations, we first derive the average position (mean valence; mean arousal) of each emotion concept across participants in each study (Fig 1B), which allows us to compute, using Euclidean distances, how misaligned each participant's placement of an emotion concept is from the typical, normative position. Averaging these distances across all concepts for each participant quantifies the overall misalignment of one's emotion-concept representations: Smaller distances indicate more aligned representations while larger distances indicate more misaligned representations. Depending on the study, participants then played one of three economic games to evaluate their social behaviors and filled out a series of questionnaires to measure their health in affective processing, enabling us to link emotion concept knowledge to both (mal)adaptive social behaviors and affective responding.

Results

We find a great deal of variability in participants' alignment to the average emotion knowledge representation of the sample (Fig 1C). While some participants' representations of emotion concepts align closely with the population average, others deviate considerably from the norm (e.g. some conceptualize 'nervous' as a negative and low arousal in contrast to the more normative sense that 'nervous' is negative and high arousal). What are the consequences of such misalignment?

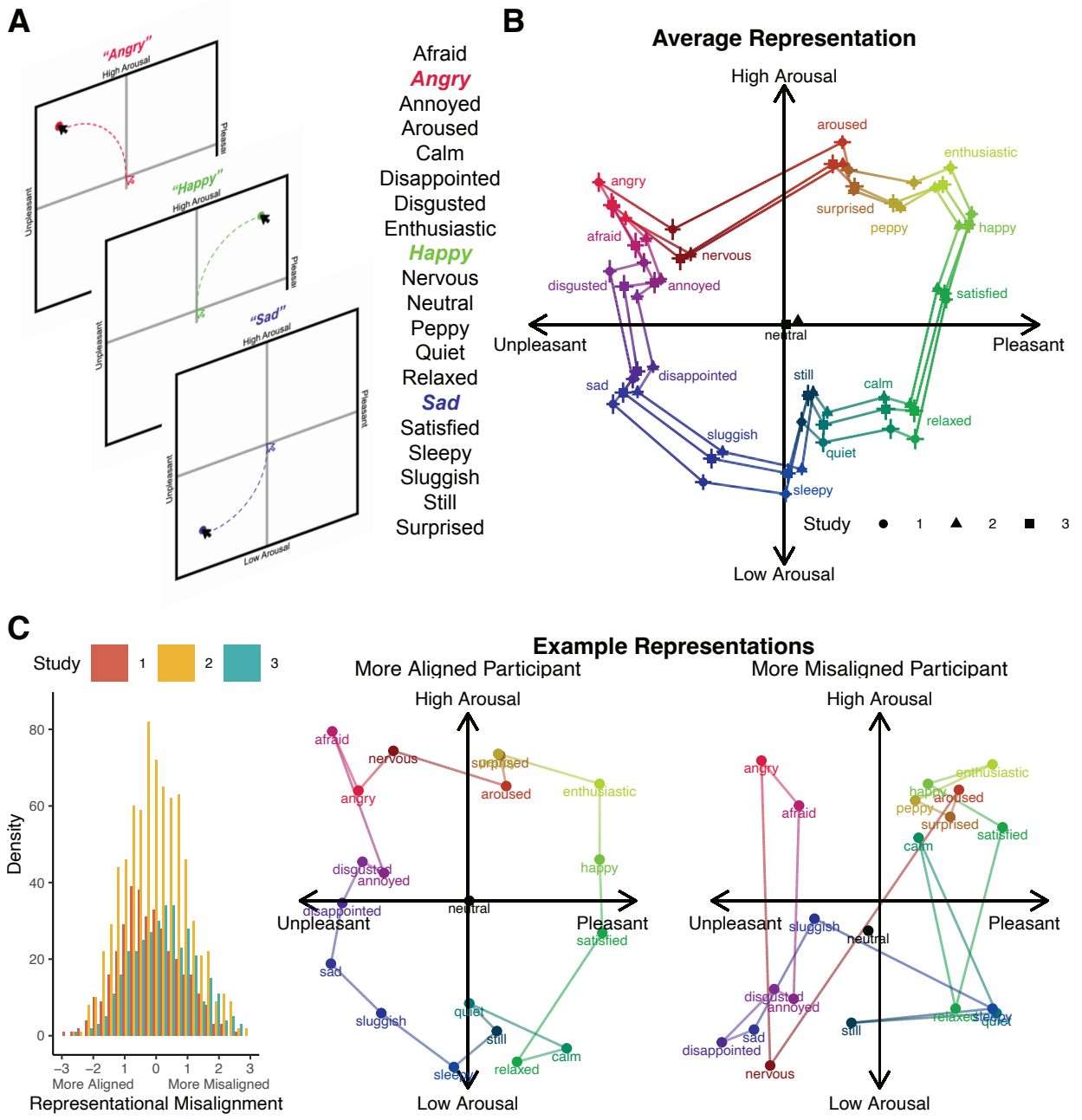

FIG 1 | Characterizing representational alignment of emotion concepts. ***(A)*** *The Emotion Classification Task. Participants are asked to draw on their emotion knowledge to place each specific emotion labels (e.g., angry) on a two-dimensional grid based on their perception of the emotion label's valence (x-axis) and arousal (y-axis).* ***(B)*** *Normative emotion representations. Depiction of the average positions of emotion labels for each of three studies. Shapes indicate the study, and the error bars indicate +/- 1 standard error above and below the average valence and arousal.* ***(C).*** *Variability in representational misalignment. Histograms depicting the distribution of individual differences in representational misalignment (log-transformed and z-scored) are plotted for each study separately (Left panel). Example representations from two participants in Study 1 that are either aligned to the group's norm or misaligned (Right panel).*

We next probe various forms of social behavior. In Studies 1 and 3, we captured sensitivity to social norms around inequality; participants decided whether to punish different partners in an Ultimatum Game where partners varied in how fairly they distributed a common pool of money amongst themselves and the participant (e.g., 50/50 to 5/95 splits). In Study 2, we measured sensitivity to social norms involving reciprocity; participants were tasked with deciding how much of a $1 endowment to contribute to a common pot in either a Prisoner's Dilemma or Public Goods Game. In both the Prisoner's Dilemma and Public Goods Game, participants possessed a unique advantage of knowing *a priori* how much their partners contributed to the common pot, which set the tone for reciprocating. To investigate how misaligned emotion knowledge shapes affective processing more generally, we administered a series of questionnaires to measure people's tendency to regulate their emotions (ERQ[47]: Study 1), as well as their symptoms of alexithymia (TAS-20[48]: Study 1-3), anhedonia (TEPS[49]: Study 2 and 3; SHAPS[50]: Study 3), apathy (AES[51]: Study 3), amotivation (BIS-BAS[52]: Study 3), and depression (CES-D[53]: Study 2 and 3). To further improve the generalizability of our findings regarding affective dysfunction, Study 3 specifically recruited participants assessed to be at risk of depression (based on CES-D scores ≥ 16). Finally, to demonstrate the specificity of representational alignment as a key driver of emotion knowledge's influence on socio-affective functioning, we control for emotion differentiation at the individual-level, which we quantify as the average pairwise Euclidean distance of participants' placement of all emotion concepts, across all analyses.

Misalignment is linked to altered social behaviors

Does a shared representation of emotion knowledge facilitate normative social behaviors? We begin by examining how representational misalignment shapes decisions to punish those who violate fairness norms in the Ultimatum Game ($N_{Study\ 1}$ = 363; $N_{Study\ 3}$ = 351). Punishment in the Ultimatum Game requires participants to pay money to prevent their partners from profiting off unfair interactions. We find that the greater the inequality in outcomes between participant and partner, the more likely a participant was to punish the partner (Study 1: b = 4.325, SE = 0.265, z = 16.342, p < .001, r = 0.512; Fig 2A. Study 3: b = 1.700, SE = 0.097, z = 17.471, p < .001, r = 0.424; Fig 2D).

This sensitivity to unfair behavior was modulated by how misaligned one's representation of emotion concepts was to the normative representation, such that the greater the misalignment, the less sensitive a participant was to how unfairly their partner behaves (Study 1: interaction b = -0.373, SE = 0.089, z = −4.196, p < .001, r = −0.102, Fig 2B; Study 3: interaction b = −0.411, SE = 0.063, z = −6.490, p < .001, r = −0.113, Fig 2E). Specifically, the more an individual's representation of emotion concepts deviated from the group representation—i.e., more misaligned, the more heavily they punished being treated unfairly, so much so that participants are willing to sacrifice personal gains to punish partners who commit relatively minor violations—far more often than participants with more aligned representations (simple effect of misalignment at inequality of 0.2, Study 1: b = 3.353, SE = 0.871, z = 3.851, p < .001, Fig 2C; Study 3: b = 3.784, SE = 0.708, z = 5.347, p < .001, Fig 2F). This suggests that representational alignment of emotion concepts is a key factor in being able to calibrate one's behavioral responses to the norm violators, and that increasingly misaligned representations is associated with greater intolerance of those who even slightly violate social norms.

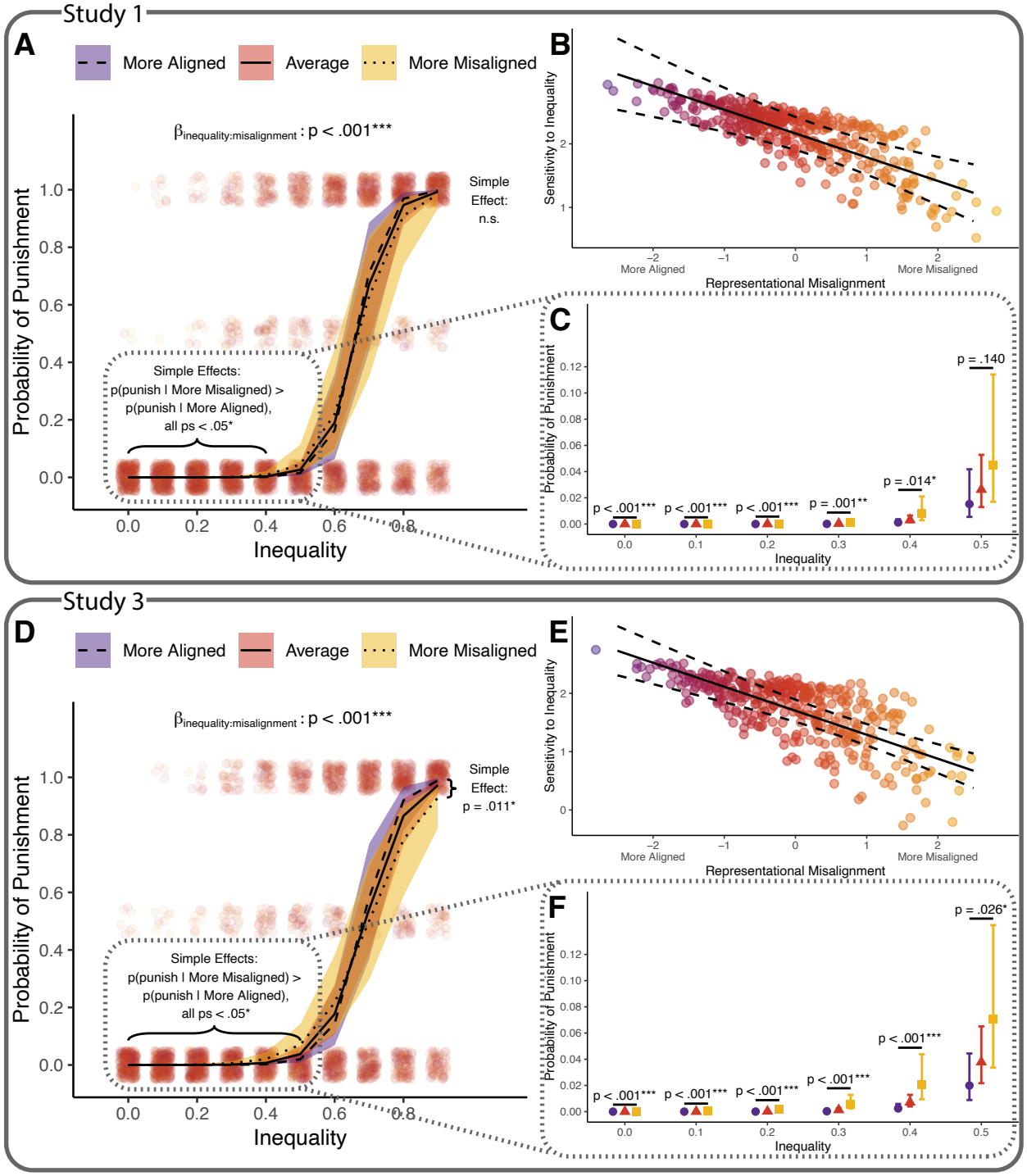

FIG 2 | Representation misalignment in emotion concepts alters likelihood of punishing norm violators. (A) & (D) *Punishment of partners who increasingly violate fairness norms is moderated by misaligned representations of emotion concepts. Each point represents the probability a participant punished a partner (playing as the Proposer) in the Ultimatum Game who acted unfairly, colored by the participants' representational misalignment. Warmer colors indicate greater misalignment, while cooler colors indicate greater alignment. Lines and shaded regions indicate the model predicted mean probability and 95% CI of punishment across levels of*

*inequality for participants with more aligned (purple: mean – 1SD), average (red), and more misaligned (yellow: mean + 1SD) representations. **(B) & (E)** Misaligned emotion-concept representations diminish sensitivity to inequality. Each point represents a participant's random coefficient (slope) for unfair outcomes as a function of their representational misalignment. **(C) & (F)** Greater representational misalignment in emotion concepts correlates with greater intolerance of even minor infractions. Error bars indicate the model-predicted mean probability and 95% CI of punishment across levels of inequality for participants with more aligned (purple: mean – 1SD), average (red), and more misaligned (yellow: mean + 1SD) representations. *p < .05, **p < .01, ***p < .001*

We next seek evidence that the link between misalignment and social behavior extends beyond the norms of inequality and fairness. We therefore measure representational misalignment alongside norms of reciprocity, in both the Prisoner's Dilemma (PD: $N_{Study\ 2A} = 306$) and Public Goods Game (PGG: $N_{Study\ 2B} = 470$). Given that these games are one-shot and played with different partners on each round, the response that produces the greatest monetary outcome is to contribute nothing to the common pot and keep all of the endowment (which is inconsistent with social norms of reciprocity). Reciprocity instead dictates that participants should match their partners' contributions.

We find that when given information about how much their partners contributed, participants tend to reciprocate in kind (PD: b = 0.644, SE = 0.023, t(303) = 27.537, p < .001, r = 0.845; PGG: b = 0.650, SE = 0.017, t(467) = 37.662, p < .001, r = 0.867; Fig 3A). This effect, however, was significantly modulated by the degree of misalignment in participants' emotion-concept representations: the greater the misalignment, the more participants contributed to the common pot even after their partners defected (PD: interaction b = −0.116, SE = 0.023, t(303) = −5.137, p < .001, r = 0.283; PGG: interaction b = −0.098, SE = 0.018, t(467) = −5.450, p < .001, r = 0.245; Fig 3B). Put differently, the greater the misalignment, the less sensitive one is to their partners' antisocial behavior. When compared to those with more aligned representations, misalignment is associated with continued, unchecked contributions to the common pot, a form of blind altruism (simple effect of misalignment when partners contributed $0: PD b = 0.178, SE = 0.029, t(303) = 6.114, p < .001, r = 0.331; PGG b = 0.176, SE = 0.028, t(467) = 6.364, p < .001, r = 0.282). Thus, across different economic game evoking distinct social norms in Studies 1 to 3, we find that representational misalignment robustly correlates with attenuated sensitivity to others' behavior that violates normative expectations.

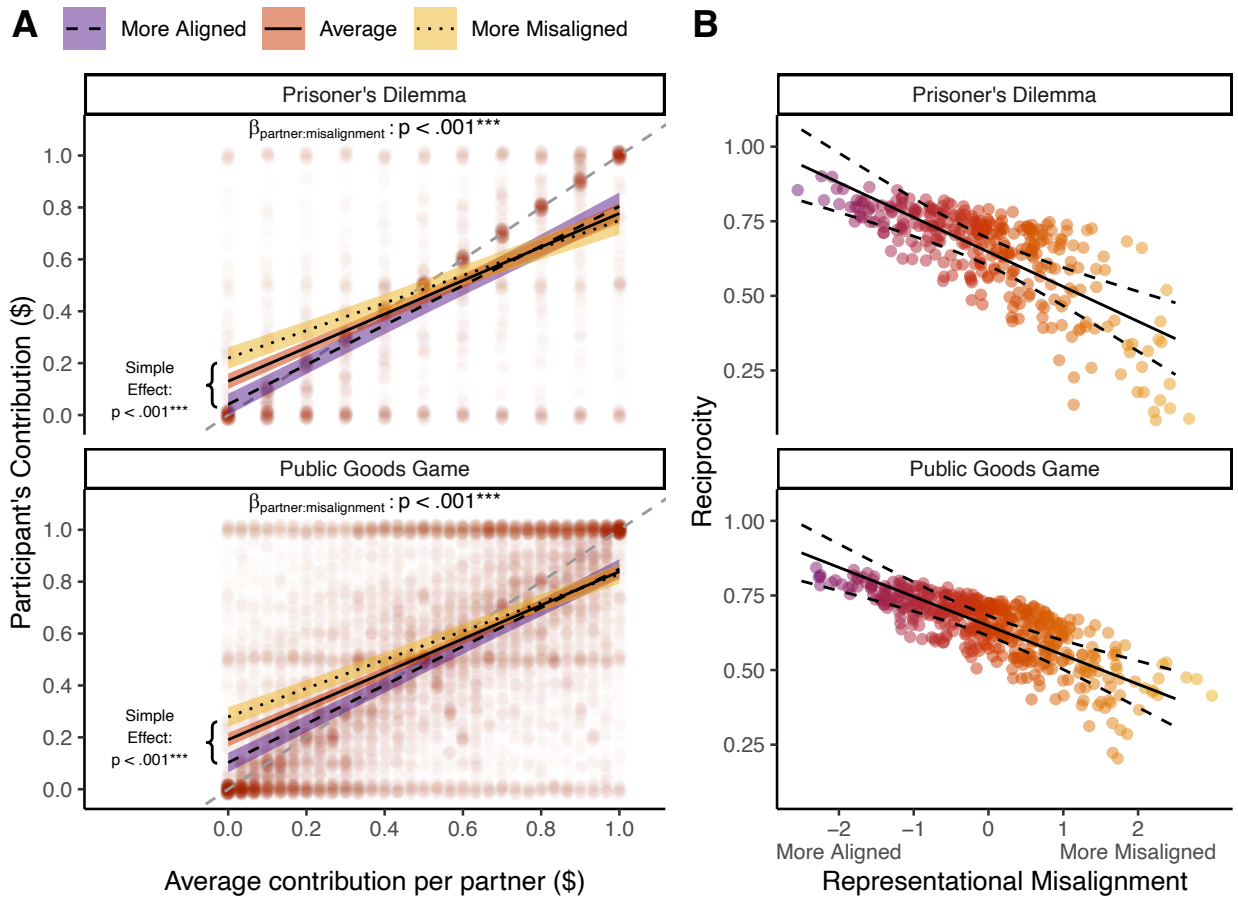

FIG 3 | Representation misalignment in emotion concepts alters social reciprocity. (A) *Reciprocity is modulated by representational misalignment of emotion concepts. Each point represents a participant's monetary contributions to a common pot as a function of their partners' contributions in either a Prisoner's Dilemma or Public Goods Game, colored by the participants' representational misalignment. Warmer colors indicate more misalignment. Cooler colors indicate more alignment. Lines and shaded regions indicate the model predicted mean and 95% CI of participant's contributions as a function of their partners' contributions for participants with more aligned (purple: mean – 1SD), average (red), and more misaligned (yellow: mean + 1SD) representations.* ***(B)*** *Misaligned emotion-concept representations diminish sensitivity to partners' social contributions. Each point represents a participant's random coefficient (slope) on their partners' contributions as a function of their representational misalignment. *p < .05, **p < .01, ***p < .001*

Misalignment predicts affective dysfunction

A key component of our theory posits that the consequences of representational misalignment in emotion concepts impact more than just social behaviors. Because people and their emotions are necessarily embedded within a broader social context, we hypothesize that misaligned representations of emotion concepts inevitably derail more general affective processing as emotions unfold in the real-world, leading to more generalized affective dysfunction. In considering which affective processes might be impaired by representational misalignment, we began by examining whether representational misalignment predicts deficits in the mobilization

of emotion knowledge. Alexithymia[48] in particular is a subclinical construct characterized by deficits in applying emotion knowledge to interpret affective experiences—i.e., difficulty identifying subjective feelings, difficulty describing subjective feelings, and diminished sensitivity to internal states—and has been closely linked to poor emotion differentiation[54–57] and impaired social behavior[58]. However, whether representational misalignment disrupts the deployment of emotion concepts remains unknown. To test this relationship, we used degree of misalignment to predict alexithymia scores (using the TAS-20)[48]. We find evidence across all three studies that greater representational misalignment is associated with more severe alexithymia (Study 1: b = 0.257, SE = 0.695, t(360) = 3.701, p < .001, r = 0.191; Study 2: b = 2.805, SE = 0.478, t(773) = 5.872, p < .001, r = 0.207; Study 3: b = 6.362, SE = 0.663, t(348) = 9.596, p < .001, r = 0.457; Fig 4). To ensure that this does not simply reflect participants' unwillingness to engage with their emotions, we also test whether representational misalignment reduces people's tendency to regulate their emotions in Study 1 (ERQ)[47] and find no evidence for this (Reappraisal: b = 0.033, SE = 0.074, t(360) = 0.446, p = .656, r = 0.024; Suppression: b = 0.173, SE = 0.102, t(360) = 1.702, p = .090, r = 0.089).

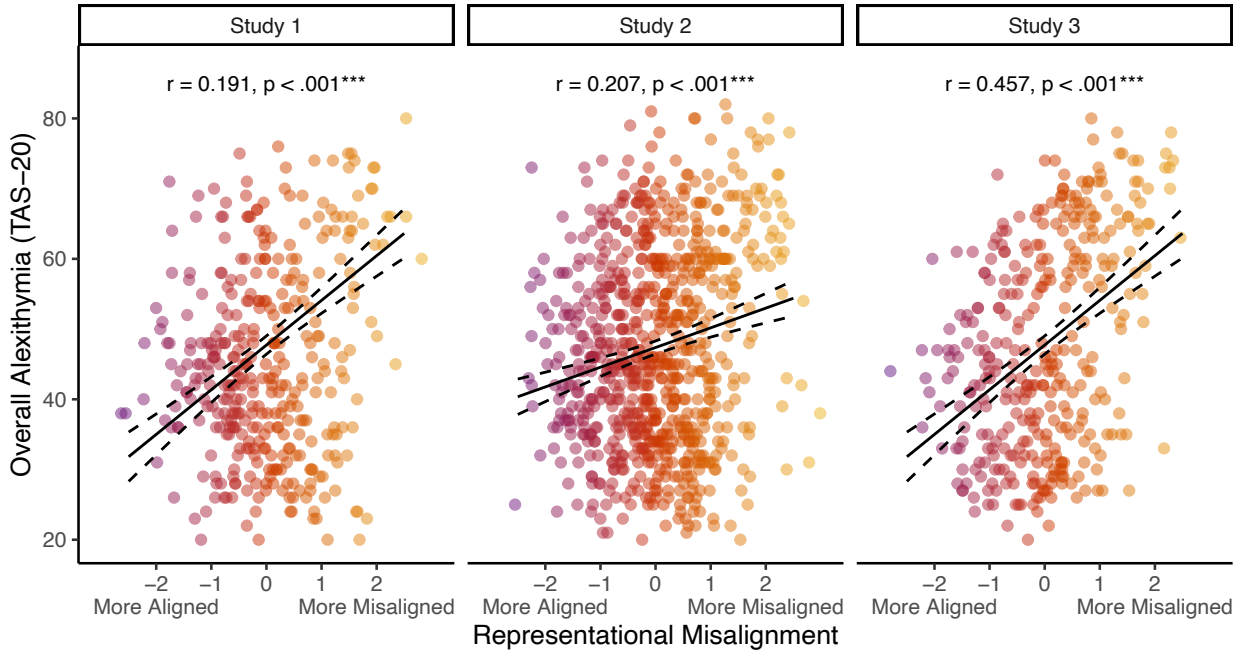

FIG 4 | Greater misalignment of emotion-concept representations correlates with more severe alexithymia. *Each point indicates a participants' total score on the Toronto Alexithymia Scale (TAS-20) as function of their representational misalignment in emotion concepts. Warmer colors indicate greater misalignment, while cooler colors indicate greater alignment. The solid line and area demarcated by the two dashed lines indicate the model-predicted mean alexithymia score and 95% CI for a given degree of misalignment. *p < .05, **p < .01, ***p < .001*

Having shown that people who possess more misaligned representations exhibit greater difficulty interpreting their affective experiences, we turn to examining the affective dysfunction seen in mood disorders[49,53]. Depression and anhedonia (i.e., the inability to experience pleasure) are strongly linked to an individual's well-being and provide additional testbeds for understanding the consequences of misaligned emotion representations. We find that representational

misalignment correlates with the severity of depressive symptoms (b = 1.840, SE = 0.449, t(773) = 4.101, p < .001, r = 0.146; Fig 5A), illustrating that deviations in one's conceptual representation of emotions from the group norm produces pervasive deficits in affective processing. Moreover, we find that representational misalignment correlates with consummatory, but not anticipatory, anhedonia (consummatory b = 0.066, SE = 0.029, t(773) = 2.264, p = .024, r = 0.081; Fig 5B; anticipatory b = −0.053, SE = 0.029, t(773) = −1.825, p = .068, r = −0.066). This pattern of results hints at these affective deficits being largely characterized by blunted experiences of pleasure, rather than difficulties anticipating pleasurable experiences. To further clarify the specificity of representational misalignment's effects on anhedonia, we leverage the expanded battery of measures in Study 3 to systematically capture and delineate the distinct, but related, affective deficits often associated with anhedonia. For example, one of the primary deficits include a diminished capacity to experience pleasure. To reduce the number of dimensions underlying these overlapping measures, we leverage exploratory factor analysis (see Methods), which identifies two latent components that distinctly map onto experiential anhedonia and intrinsic motivation (Table 1).

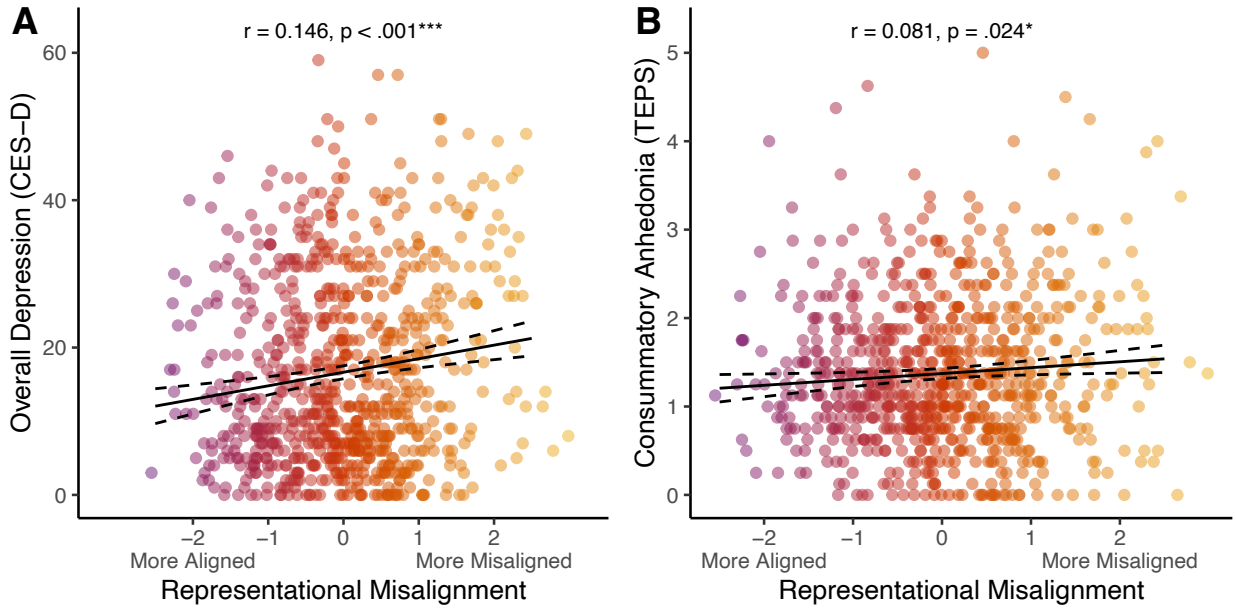

FIG 5 | Greater misalignment of emotion-concept representations correlates with affective dysfunction. ***(A)*** *Representational misalignment is associated with more severe depression in Study 2. Each point indicates a participant's total score on the Center for Epidemiologic Studies Depression Scale (CES-D) as function of their representational misalignment in emotion concepts.* ***(B)*** *Representational misalignment is associated with more severe anhedonia in Study 1. Each point indicates a participant's score on the Consummatory subscale of the Temporal Experience of Pleasure Scale (TEPS) as function of their representational misalignment in emotion concepts. Warmer colors indicate greater misalignment, while cooler colors indicate greater alignment. The solid line and area demarcated by the two dashed lines indicate the model-predicted score and 95% CI for a given degree of misalignment. *p < .05, **p < .01, ***p < .001*

TABLE 1 | Exploratory factor analysis identifies two latent factors underlying measures associated with anhedonia

Measure	Factor Loading	
	Experiential Anhedonia	Intrinsic Motivation
Snaith-Hamilton Pleasure Scale (SHAPS)	-0.942*	-0.104
Temporal Experience of Pleasure Scale (TEPS: Anticipatory Subscale)	-0.572*	0.270
Temporal Experience of Pleasure Scale (TEPS: Consummatory Subscale)	-0.595*	0.163
Apathy Evaluation Scale (AES)	0.863*	0.085
Behavioral Activaton Scale (BAS: Reward Responsiveness Subscale)	-0.563*	0.241
Behavioral Activaton Scale (BAS: Drive Subscale)	0.047	0.702*
Behavioral Activaton Scale (BAS: Fun-Seeking Subscale)	0.165	0.860*
Behavioral Inhibition Scale (BIS)	-0.221	-0.098

*Note: *Loading ≥ 0.50 which indicates that this (sub)scale loaded reliably onto this latent factor and was used in structural equation modelling to define the latent factor.*

Using structural equation modelling (SEM), we find that greater representational misalignment is robustly associated with depression severity (total effect b = 0.044, SE = 0.006, z = 7.391, p < .001, r = 0.332) and the latent factor of experiential anhedonia (total effect b = 0.008, SE = 0.003, z = 2.339, p = .019, r = 0.129). SEM additionally reveals that the blunting of pleasurable experiences due to misalignment is primarily mediated through alexithymia (indirect effect b = 0.010, SE = 0.002, z = 5.301, p < .001, r = 0.179; direct effect b = -0.003, SE = 0.003, z = -0.848, p = .396, r = -0.049; Fig 6). In other words, individuals with misaligned representations of emotion concepts experience diminished pleasure in large part because they exhibit deficits in the deployment of their emotion-concept knowledge. This in turn partially explains the severity of depressive symptoms (indirect serial mediation effect: b = 0.008, SE = 0.002, z = 3.300, p = .001, r = 0.063). Beyond these consequences of anhedonia, alexithymia also independently explains part of the severity in participants' depressive symptoms (b = 0.309, SE = 0.056, z = 5.477, p < .001, r = 0.350), indicating that the challenges associated with deploying misaligned representations of emotion concept might uniquely exacerbate depressive symptoms (indirect effect: b = 0.019, SE = 0.004, z = 4.779, p < .001, r = 0.146).

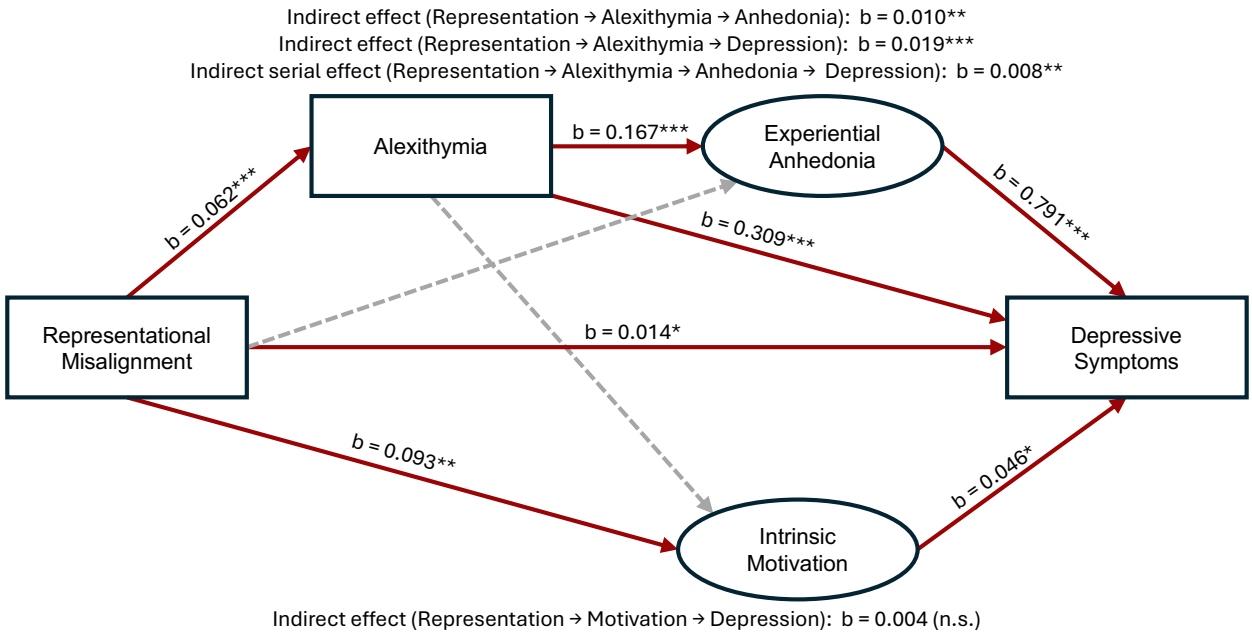

FIG 6 | Structural equation model of representational misalignment and affective dysfunction. *Graph indicates the associations between variables of the structural equation model. Rectangular borders identify variables that are directly measured. Elliptical borders identify latent variables comprising multiple measurements. Solid arrows indicate significant positive associations between variables. Dashed arrows indicate non-significant associations. *p < .05, **p < .01, ***p < .001*

Surprisingly, we find that representational misalignment correlates not with a lack of motivation, but instead with greater intrinsic motivation (total effect: b = 0.105, SE = 0.031, z = 3.335, p = 0.001, r = 0.208). That is, individuals with more misaligned representations of emotion concepts are highly driven to pursue self-generated goals. Greater intrinsic motivation also correlates with more severe depressive symptoms (b = 0.046, SE = 0.021, z = 2.190, p = .029, r = 0.176), although this did not explain the link between misaligned emotion representations and depression (indirect effect: b = 0.004, SE = 0.002, z = 1.712, p = .087, r = 0.032). Intrinsic motivation was also unrelated to alexithymia (b = 0.190, SE = 0.268, z = 0.708, p = .479, r = 0.056), suggesting that the association between representational alignment and intrinsic motivation is unlikely to be driven by one's deployment of their idiosyncratic emotion knowledge. Regardless, even after accounting for all indirect effects, misaligned emotion representations continue to explain unique variability in the severity of participants' depressive symptoms (b = 0.014, SE = 0.006, z = 2.322, p = .020, r = 0.104).

Discussion

People's conceptual knowledge about emotions critically shapes how they experience[10,24,25,59,60] and regulate their affective responses to emotionally evocative events[37,61–63]. Research has primarily focused on individual differences in emotion differentiation and the consequences of poor differentiation[28–36]. This has neglected the fundamentally social ways in which people acquire and use their emotion knowledge. This key gap in understanding how emotion knowledge shapes adaptive behavior suggests there may be an overlooked mechanism

that helps to explain maladaptive social behaviors and emotional processing. Across three studies with nearly 1,500 participants, we close this gap and interrogate how deviations in emotion-concept representations from the group norm undermines social and emotional well-being.

By characterizing the substantial variability in people's representations of emotion concepts and their deviations from the population-level representation, we find that greater representational misalignment of emotion concepts reduces sensitivity to pertinent social information. This blunted sensitivity predicts less strategic punishment of norm violators and poorer coordination with others across three different social contexts. Individuals with more misaligned representations tended to be less attuned or judicious when responding to unfair behavior. Those with the greatest misalignment were highly intolerant of even the slightest infraction, such that they consistently punished others for minor norm violations even when it came at great personal cost to themselves. This diminished sensitivity to social information extends beyond just norms of inequality. We also observed that individuals with more misaligned representations also adhered less to norms surrounding reciprocity, contributing more to a common pot even when armed with knowledge of their partners' defections. This inflexibility in behavioral responding resulted in poorer personal financial outcomes. Together, these findings indicate that the degree to which an individual shares conceptual knowledge of emotions with the broader group plays a key role in shaping socially well-oiled behaviors[11,14]. Indeed, aligning one's emotion-concept representations to the social consensus may be necessary to calibrate one's actions when responding to others[27,41,42].

Furthermore, in acknowledging that affective processes necessarily play out in the context of people's interactions and relationships with others in the real world[41,44,45,64,65], our results expand previously limiting conceptions of emotion knowledge's role in affective processes. We find evidence that social dimensions of emotion knowledge have powerful implications for emotional well-being more generally. Representational misalignment in emotion concepts correlate with persistent affective dysfunction including more severe alexithymia, anhedonia, and depression. More misaligned emotion-concept representations hindered people's ability to deploy this knowledge to interpret their own experiences, which in turn blunted their experiences of pleasure and exacerbated depressive symptoms. Building on an emerging body of work documenting the ubiquity of socio-emotional interactions where people seek others out for help regulating their own emotions[66,67], our results suggest that an individual's personal and emotional well-being is inextricable from their ability to adaptively navigate their social interactions and relationships, and that shared representations of emotion concepts play a central role in mediating these socio-affective processes.

Yet, despite these new insights, much remains unknown about how exactly emotion knowledge shapes social cognition. Our work here lays the foundation to uncover the precise mechanisms underlying the role emotion knowledge plays in shaping the dynamics of socio-emotional interactions. One obvious area for future work is understanding whether and how alignment in emotion-concept representations support empathic accuracy in social interactions[68,69]. We suspect that people who share a common understanding of emotion concepts likely communicate more efficiently and effectively with each other, facilitating more accurate perception and interpretation of the other's mental states and affective experiences[43,70–72]. At the same time, empathic accuracy serves as a necessary precondition for effective interpersonal emotion regulation[73] but it is unclear whether those with more misaligned representations of emotion concepts benefit less from others' efforts to regulate their emotions.

This raises another interesting avenue for future work. How do disruptions of these fundamental socio-emotional processes due to misaligned representations of emotion concepts shape people's ability to develop close and meaningful social connections? Given that much work documents how empathic accuracy is a key determinant of satisfaction in long-term romantic relationships[7,71,73], it seems apparent that misaligned representations of emotion concepts would undermine relationship satisfaction. However, whether misaligned emotion-representations hinder the formation of these relationships in the first place remain wholly unknown. Understanding representational misalignment's cascading effects across this suite of socio-emotional processes would reveal fundamental insights that might help explain the persistent co-occurrence of social and affective dysfunction[74–78], and potentially inspire more effective interventions that improve people's social and emotional well-being.

Methods

Participants

Across three studies, participants (N = 1,820) completed an emotion classification task followed by one of three economic games and a set of questionnaires. In Study 1 (N = 398), participants completed a series Ultimatum Games, the Toronto Alexithymia Scale (TAS-20)[48] and the Emotion Regulation Questionnaire (ERQ)[47]. In Study 2, participants either completed a series of Prisoner's Dilemmas (Study 2A N = 395) or Public Goods Games (Study 2B N = 519) in addition to the TAS-20, Center for Epidemiologic Studies Depression Scale (CES-D)[53], and the Temporal Experience of Pleasure Scale (TEPS)[49]. In Study 3 (N = 508), participants completed a series of Ultimatum Games, the TAS-20, CES-D, TEPS, as well as the Snaith-Hamilton Pleasure Scale (SHAPS)[50], Apathy Evaluation Scale (AES)[51], and Behavioral Inhibition/Activation Scale (BIS-BAS)[52]. Using preregistered criteria based on existing work[27,46], we excluded participants (N=330 total; Study 1 exclusion = 34; Study 2A exclusion = 89; Study 2B exclusion = 49; Study 3 exclusion = 157) who placed the "neutral" label outside of a 100 × 100-pixel square in the center (i.e., participants were instructed to place "neutral" in the center of the 500 x 500 pixelated grid). We excluded an additional participant from Study 1 because they did not complete any questionnaires.

The final sample was N = 393 for Study 1 (171 Female, 192 Male, Age: Mean = 33.8, SD = 9.85,Range = [18, 73], one participant entered their age as 3 and is not included in these descriptive statistics for age), N = 306 for Study 2A (131 Female, 175 Male, Age: Mean = 35.6, SD = 11.1,Range = [19, 73] , one participant entered their age as 4 and is not included in these descriptive statistics for age), N = 470 for Study 2B (238 Female, 232 Male, Age: Mean = 33.0, SD = 10.5,Range = [18, 75]), and N = 351 for Study 3 (211 Female, 297 Male, Age: Mean = 34.1, SD = 9.33 ,Range = [19, 73]).Thus, the final sample across all three studies was N = 1490. Participants were recruited from Amazon Mechanical Turk and received monetary compensation and provided informed consent in a manner approved by Brown University's Institutional Review Board under protocol 1607001555. The data reported across Studies 1 to 3 are part of a dataset collected for previously published studies[27,46].

Procedure

In the Emotion Classification Task, participants placed a series of emotion labels on a 500 x 500 pixelated grid with two dimensions corresponding to valence (x-axis) and arousal (y-axis). The twenty labels included in the stimulus set consisted of neutral, surprised, aroused, peppy,

enthusiastic, happy, satisfied, relaxed, calm, sleepy, still, quiet, sluggish, sad, disappointed, disgusted, annoyed, angry, afraid, and nervous. These labels were chosen on the basis of prior work suggesting that these words evenly spanned the two-dimensional space of valence and arousal[26] or are heavily implicated in normative social behavior[79].

Participants then completed either a series of Ultimatum Games , Prisoner's Dilemmas, or Public Goods Games. In Studies 1 and 3, participants completed 20 one-shot Ultimatum Games as a responder or third party. Since prior work shows that responders and third parties reacted to unfairness in the same ways, these data were collapsed. On each trial participants were asked to decide whether to punish the proposer who made one of twenty offers varying from 5/95 to 50/50. Participants in the Study 2A played 22 rounds of the Prisoner's Dilemma, each with a new partner. They were instructed that both them and their partners were given $1 each and asked how much of this endowment (in $0.10 increments) they would like to contribute to a common pot that their partner could also contribute to. The total amount in the pot would then be multiplied 1.5x and redistributed evenly to the participant and the partner. Crucially, this version of the Prisoner's Dilemma operated sequentially, such that participants always knew what their partners had already contributed and could always keep all their endowment to maximize their outcomes at their partners' expense. Participants in Study 2B played 62 rounds of the Public Goods Game which essentially constituted a multiplayer version of the Prisoner's Dilemma with three partners. Again, participants had full knowledge of their partner's contributions prior to making their decision of how much to contribute to the common pot. Partners' total contributions ranged uniformly from $0 to $3 in $0.10 increments.

Exploratory factor analysis

We conducted exploratory factor analysis to identify common factors underlying affective dysfunction associated with anhedonia. We included the scores from the full SHAPS and AES scales, and the subscales of TEPS (anticipatory and consummatory), BIS-BAS (BIS, BAS-reward responsiveness, BAS-drive, and BAS-fun-seeking) as items of the analysis. We assumed that the latent dimensions might be related and used the oblique rotation method, promax. We used the scree test to determine how many factors to extract and ultimately decided on 2 factors. The scree plot is shown in Supplementary Fig 1. The 2-factor solution explained 51.6% of the variance. After inspecting the variables that loaded at least .5 on each factor, we determined that the two factors were best labelled experiential anhedonia (proportion of variance = 33.9%) and intrinsic motivation (proportion of variance = 17.7%). The experiential anhedonia factor consisted of SHAPS, AES, both TEPS subscales, and the BAS-reward responsiveness subscale, whereas the intrinsic motivation factor consisted of the BAS-drive and BAS-fun-seeking subscales. The BIS was not retained in the final solution because it did not load strongly onto either factor.

Structural equation modelling

Based on this factor structure, we then conducted structural equation modelling to examine the associations between representational misalignment, alexithymia, depression, and our latent factors of experiential anhedonia and intrinsic motivation. Our structural equation model had good fit (SRMR = 0.034 ≤ 0.08, CFI = 0.984 ≥ 0.95, RMSEA = 0.074 ≤ 0.08, 90% CI = [0.049, 0.101]). The full model is specified in Table 1.

Statistical models

All statistical analyses were conducted in RStudio v2023.12.1+402 (RStudio Team, 2015) with R v4.3.3 (R Core Team, 2017). We fitted linear mixed effects models to continuous data with degrees of freedom estimated using the Satterthwaite method, and logistic mixed-effects model to binary data using the *lme4* package v1.1-35.2[80,81], with maximal random effects where possible. Simple slopes and comparisons were extracted using the *marginaleffects* package v0.21.0[82]. Structural equation modelling was conducted using the *lavaan* package v0.6-19[83]. Plots of participant data and model predictions were generated using the *ggplot2* v3.5.0[84]and *ggeffects* v1.5.1[85] packages. Stimuli presentation and the recording of participants' responses were supported by *Qualtrics.*

Data &Code Availability

All necessary data and code necessary to reproduce the results are openly accessible at https://osf.io/8cs5h/overview?view_only=6450d176ddfc4675a51f7f1e7b3f67ad.

References:

1. Barrett, L. F. & Bliss-Moreau, E. Affect as a Psychological Primitive. in *Advances in Experimental Social Psychology* vol. 41 167–218 (Academic Press, 2009). DOI:10.1016/S0065-2601(08)00404-8.
2. Taschereau-Dumouchel, V., Michel, M., Lau, H., Hofmann, S. G. & LeDoux, J. E. Putting the "mental" back in "mental disorders": a perspective from research on fear and anxiety. *Mol Psychiatry* **27**, 1322–1330 (2022) DOI:10.1038/s41380-021-01395-5.
3. LeDoux, J. E. & Hofmann, S. G. The subjective experience of emotion: a fearful view. *Current Opinion in Behavioral Sciences* **19**, 67–72 (2018) DOI:10.1016/j.cobeha.2017.09.011.
4. Górna, K., Szpalik, R., Rybakowski, J. K. & Jaracz, K. Health behaviours of patients with affective disorders: a cross-sectional study. *BMC Psychiatry* **23**, 561 (2023) DOI:10.1186/s12888-023-05056-5.
5. Cohen, S. & Rodriguez, M. S. Pathways linking affective disturbances and physical disorders. *Health Psychology* **14**, 374–380 (1995) DOI:10.1037/0278-6133.14.5.374.
6. Jardine, B. B., Vannier, S. & Voyer, D. Emotional intelligence and romantic relationship satisfaction: A systematic review and meta-analysis. *Personality and Individual Differences* **196**, 111713 (2022) DOI:10.1016/j.paid.2022.111713.
7. Sened, H. *et al.* Empathic accuracy and relationship satisfaction: A meta-analytic review. *Journal of Family Psychology* **31**, 742–752 (2017) DOI:10.1037/fam0000320.
8. Mazzuca, S., Kafetsios, K., Livi, S. & Presaghi, F. Emotion regulation and satisfaction in long-term marital relationships: The role of emotional contagion. *Journal of Social and Personal Relationships* **36**, 2880–2895 (2019) DOI:10.1177/0265407518804452.
9. Jackson, J. C. *et al.* Emotion semantics show both cultural variation and universal structure. *Science* **366**, 1517–1522 (2019) DOI:10.1126/science.aaw8160.
10. Barrett, L. F. Feelings or Words? Understanding the Content in Self-Report Ratings of Experienced Emotion. *Journal of Personality and Social Psychology* **87**, 266–281 (2004) DOI:10.1037/0022-3514.87.2.266.
11. Barrett, L. F. Emotions are real. *Emotion* **12**, 413–429 (2012) DOI:10.1037/a0027555.
12. Clore, G. L. & Ortony, A. Psychological construction in the OCC model of emotion. *Emotion Review* **5**, 335–343 (2013).
13. Russell, J. A. Core affect and the psychological construction of emotion. *Psychological Review* **110**, 145–172 (2003) DOI:10.1037/0033-295X.110.1.145.
14. Lindquist, K. A., Jackson, J. C., Leshin, J., Satpute, A. B. & Gendron, M. The cultural evolution of emotion. *Nat Rev Psychol* **1**, 669–681 (2022) DOI:10.1038/s44159-022-00105-4.
15. Ekman, P. & Cordaro, D. What is Meant by Calling Emotions Basic. *Emotion Review* **3**, 364–370 (2011) DOI:10.1177/1754073911410740.
16. Ekman, P. & Davidson, R. J. *The nature of emotion: Fundamental questions.* (Oxford University Press, 1994).
17. Cowen, A. S., Elfenbein, H. A., Laukka, P. & Keltner, D. Mapping 24 emotions conveyed by brief human vocalization. *American Psychologist* **74**, 698–712 (2019) DOI:10.1037/amp0000399.
18. Cowen, A. S., Laukka, P., Elfenbein, H. A., Liu, R. & Keltner, D. The primacy of categories in the recognition of 12 emotions in speech prosody across two cultures. *Nat Hum Behav* **3**, 369–382 (2019) DOI:10.1038/s41562-019-0533-6.

19. Cowen, A. S. & Keltner, D. Self-report captures 27 distinct categories of emotion bridged by continuous gradients. *Proceedings of the National Academy of Sciences* **114**, E7900–E7909 (2017) DOI:10.1073/pnas.1702247114.
20. Cowen, A. S. & Keltner, D. Clarifying the Conceptualization, Dimensionality, and Structure of Emotion: Response to Barrett and Colleagues. *Trends in Cognitive Sciences* **22**, 274–276 (2018) DOI:10.1016/j.tics.2018.02.003.
21. Cowen, A. S. & Keltner, D. Semantic Space Theory: A Computational Approach to Emotion. *Trends in Cognitive Sciences* **25**, 124–136 (2021) DOI:10.1016/j.tics.2020.11.004.
22. Keltner, D., Brooks, J. A. & Cowen, A. Semantic Space Theory: Data-Driven Insights Into Basic Emotions. *Curr Dir Psychol Sci* **32**, 242–249 (2023) DOI:10.1177/09637214221150511.
23. Horikawa, T., Cowen, A. S., Keltner, D. & Kamitani, Y. The Neural Representation of Visually Evoked Emotion Is High-Dimensional, Categorical, and Distributed across Transmodal Brain Regions. *iScience* **23**, (2020) DOI:10.1016/j.isci.2020.101060.
24. Barrett, L. F. Discrete Emotions or Dimensions? The Role of Valence Focus and Arousal Focus. *Cognition and Emotion* **12**, 579–599 (1998) DOI:10.1080/026999398379574.
25. Feldman, L. A. Valence focus and arousal focus: Individual differences in the structure of affective experience. *Journal of Personality and Social Psychology* **69**, 153–166 (1995) DOI:10.1037/0022-3514.69.1.153.
26. Feldman, L. A. Variations in the Circumplex Structure of Mood. *Pers Soc Psychol Bull* **21**, 806–817 (1995) DOI:10.1177/0146167295218003.
27. Heffner, J. & FeldmanHall, O. A probabilistic map of emotional experiences during competitive social interactions. *Nat Commun* **13**, 1718 (2022) DOI:10.1038/s41467-022-29372-8.
28. Nook, E. C. Emotion Differentiation and Youth Mental Health: Current Understanding and Open Questions. *Front. Psychol.* **12**, (2021) DOI:10.3389/fpsyg.2021.700298.
29. Smidt, K. E. & Suvak, M. K. A brief, but nuanced, review of emotional granularity and emotion differentiation research. *Current Opinion in Psychology* **3**, 48–51 (2015) DOI:10.1016/j.copsyc.2015.02.007.
30. Kashdan, T. B., Barrett, L. F. & McKnight, P. E. Unpacking Emotion Differentiation: Transforming Unpleasant Experience by Perceiving Distinctions in Negativity. *Curr Dir Psychol Sci* **24**, 10–16 (2015) DOI:10.1177/0963721414550708.
31. Thompson, R. J., Springstein, T. & Boden, M. Gaining clarity about emotion differentiation. *Social and Personality Psychology Compass* **15**, e12584 (2021) DOI:10.1111/spc3.12584.
32. DeLap, G. A. L., Vine, V., Santee, A. C. & Starr, L. R. Putting it into words: Emotion vocabulary, emotion differentiation, and depression among adolescents. *Emotion* **25**, 102–113 (2025) DOI:10.1037/emo0001429.
33. Demiralp, E. *et al.* Feeling Blue or Turquoise? Emotional Differentiation in Major Depressive Disorder. *Psychol Sci* **23**, 1410–1416 (2012) DOI:10.1177/0956797612444903.
34. Liu, D. Y., Gilbert, K. E. & Thompson, R. J. Emotion differentiation moderates the effects of rumination on depression: A longitudinal study. *Emotion* **20**, 1234–1243 (2020) DOI:10.1037/emo0000627.
35. Ottenstein, C. Emotion regulation effectiveness accounts for the associations of self-reported emotion differentiation with well-being and depression. *Cognition and Emotion* **34**, 994–1002 (2020) DOI:10.1080/02699931.2019.1691506.

36. Starr, L. R., Hershenberg, R., Li, Y. I. & Shaw, Z. A. When Feelings Lack Precision: Low Positive and Negative Emotion Differentiation and Depressive Symptoms in Daily Life. *Clinical Psychological Science* **5**, 613–631 (2017) DOI:10.1177/2167702617694657.
37. Lindquist, K. A. Language is Powerful. *Emotion Review* **1**, 16–18 (2009) DOI:10.1177/1754073908097177.
38. Fedorenko, E., Piantadosi, S. T. & Gibson, E. A. F. Language is primarily a tool for communication rather than thought. *Nature* **630**, 575–586 (2024) DOI:10.1038/s41586-024-07522-w.
39. Wierzbicka, A. Language and Metalanguage: Key Issues in Emotion Research. *Emotion Review* **1**, 3–14 (2009) DOI:10.1177/1754073908097175.
40. Gallois, C. The Language and Communication of Emotion: Universal, Interpersonal, or Intergroup? *American Behavioral Scientist* **36**, 309–338 (1993) DOI:10.1177/0002764293036003005.
41. Van Kleef, G. A. How emotions regulate social life: The emotions as social information (EASI) model. *Current Directions in Psychological Science* **18**, 184–188 (2009).
42. Hutcherson, C. A. & Gross, J. J. The moral emotions: A social–functionalist account of anger, disgust, and contempt. *Journal of personality and social psychology* **100**, 719 (2011).
43. Hall, J. A. & Schmid Mast, M. Sources of accuracy in the empathic accuracy paradigm. *Emotion* **7**, 438–446 (2007) DOI:10.1037/1528-3542.7.2.438.
44. Butler, E. A. Temporal Interpersonal Emotion Systems: The "TIES" That Form Relationships. *Pers Soc Psychol Rev* **15**, 367–393 (2011) DOI:10.1177/1088868311411164.
45. Butler, E. A. Emotions are temporal interpersonal systems. *Current Opinion in Psychology* **17**, 129–134 (2017) DOI:10.1016/j.copsyc.2017.07.005.
46. Heffner, J., Son, J.-Y. & FeldmanHall, O. Emotion prediction errors guide socially adaptive behaviour. *Nature Human Behaviour* **5**, 1391–1401 (2021) DOI:10.1038/s41562-021-01213-6.
47. Gross, J. J. & John, O. P. Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *Journal of personality and social psychology* **85**, 348 (2003).
48. Bagby, R. M., Parker, J. D. A. & Taylor, G. J. Twenty-five years with the 20-item Toronto Alexithymia Scale. *Journal of Psychosomatic Research* **131**, 109940 (2020) DOI:10.1016/j.jpsychores.2020.109940.
49. Gard, D. E., Gard, M. G., Kring, A. M. & John, O. P. Anticipatory and consummatory components of the experience of pleasure: a scale development study. *Journal of research in personality* **40**, 1086–1102 (2006).
50. Franken, I. H. A., Rassin, E. & Muris, P. The assessment of anhedonia in clinical and non-clinical populations: Further validation of the Snaith–Hamilton Pleasure Scale (SHAPS). *Journal of Affective Disorders* **99**, 83–89 (2007) DOI:10.1016/j.jad.2006.08.020.
51. Marin, R. S., Biedrzycki, R. C. & Firinciogullari, S. Reliability and validity of the apathy evaluation scale. *Psychiatry Research* **38**, 143–162 (1991) DOI:10.1016/0165-1781(91)90040-V.
52. Carver, C. S. & White, T. L. Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology* **67**, 319–333 (1994) DOI:10.1037/0022-3514.67.2.319.
53. Radloff, L. S. The CES-D scale: A self-report depression scale for research in the general population. *Applied psychological measurement* **1**, 385–401 (1977).

54. Edwards, E. R., Shivaji, S., Micek, A. & Wupperman, P. Distinguishing alexithymia and emotion differentiation conceptualizations through linguistic analysis. *Personality and Individual Differences* **157**, 109801 (2020) DOI:10.1016/j.paid.2019.109801.
55. Erbas, Y., Ceulemans, E., Lee Pe, M., Koval, P. & Kuppens, P. Negative emotion differentiation: Its personality and well-being correlates and a comparison of different assessment methods. *Cognition and Emotion* **28**, 1196–1213 (2014) DOI:10.1080/02699931.2013.875890.
56. Erbas, Y. *et al.* Emotion differentiation dissected: between-category, within-category, and integral emotion differentiation, and their relation to well-being. *Cognition and Emotion* **33**, 258–271 (2019) DOI:10.1080/02699931.2018.1465894.
57. Luminet, O. & Nielson, K. A. Alexithymia: Toward an Experimental, Processual Affective Science with Effective Interventions. *Annual Review of Psychology* **76**, 741–769 (2025) DOI:10.1146/annurev-psych-021424-030718.
58. FeldmanHall, O., Dalgleish, T. & Mobbs, D. Alexithymia decreases altruism in real social decisions. *Cortex* **49**, 899–904 (2013) DOI:10.1016/j.cortex.2012.10.015.
59. Barrett, L. F., Lindquist, K. A. & Gendron, M. Language as context for the perception of emotion. *Trends in Cognitive Sciences* **11**, 327–332 (2007) DOI:10.1016/j.tics.2007.06.003.
60. Barrett, L. F. *et al.* The Theory of Constructed Emotion: More Than a Feeling. *Perspect Psychol Sci* **20**, 392–420 (2025) DOI:10.1177/17456916251319045.
61. Izard, C. E. *et al.* Emotion Knowledge, Emotion Utilization, and Emotion Regulation. *Emotion Review* **3**, 44–52 (2011) DOI:10.1177/1754073910380972.
62. Barrett, L. F., Gross ,James, Christensen ,Tamlin Conner & and Benvenuto, M. Knowing what you're feeling and knowing what to do about it: Mapping the relation between emotion differentiation and emotion regulation. *Cognition and Emotion* **15**, 713–724 (2001) DOI:10.1080/02699930143000239.
63. Tamir, M. *et al.* Desired emotions across cultures: A value-based account. *Journal of Personality and Social Psychology* **111**, 67–82 (2016) DOI:10.1037/pspp0000072.
64. Van Kleef, G. A. The emerging view of emotion as social information. *Social and Personality Psychology Compass* **4**, 331–343 (2010).
65. Van Kleef, G. A. & Côté, S. The Social Effects of Emotions. *Annu. Rev. Psychol.* **73**, 629–658 (2022) DOI:10.1146/annurev-psych-020821-010855.
66. Williams, W. C., Morelli, S. A., Ong, D. C. & Zaki, J. Interpersonal emotion regulation: Implications for affiliation, perceived support, relationships, and well-being. *Journal of personality and social psychology* **115**, 224 (2018).
67. Zaki, J. & Williams, W. C. Interpersonal emotion regulation. *Emotion* **13**, 803–810 (2013) DOI:10.1037/a0033839.
68. Ickes, W. Empathic Accuracy. *Journal of Personality* **61**, 587–610 (1993) DOI:10.1111/j.1467-6494.1993.tb00783.x.
69. Ickes, W., Stinson, L., Bissonnette, V. & Garcia, S. Naturalistic social cognition: Empathic accuracy in mixed-sex dyads. *Journal of personality and social psychology* **59**, 730 (1990).
70. Zaki, J., Bolger, N. & Ochsner, K. It Takes Two: The Interpersonal Nature of Empathic Accuracy. *Psychol Sci* **19**, 399–404 (2008) DOI:10.1111/j.1467-9280.2008.02099.x.
71. Berlamont, L. *et al.* Couple similarity in empathic accuracy and relationship well-being. *Journal of Social and Personal Relationships* **41**, 1301–1323 (2024) DOI:10.1177/02654075231206412.

72. Soto, J. A. & Levenson, R. W. Emotion recognition across cultures: The influence of ethnicity on empathic accuracy and physiological linkage. *Emotion* **9**, 874–884 (2009) DOI:10.1037/a0017399.
73. Arican-Dinc, B. & Gable, S. L. Empathic accuracy and interpersonal emotion regulation in close relationships. *Emotion* **25**, 725–735 (2025) DOI:10.1037/emo0001448.
74. Gotlib, I. H. Interpersonal and cognitive aspects of depression. *Current Directions in Psychological Science* **1**, 149–154 (1992).
75. Kupferberg, A., Bicks, L. & Hasler, G. Social functioning in major depressive disorder. *Neuroscience & Biobehavioral Reviews* **69**, 313–332 (2016) DOI:10.1016/j.neubiorev.2016.07.002.
76. Rottenberg, J. & Gotlib, I. H. Socioemotional functioning in depression. *Mood disorders: A handbook of science and practice* 61–77 (2004).
77. Oakley, B. F. M. *et al.* Alexithymia in autism: cross-sectional and longitudinal associations with social-communication difficulties, anxiety and depression symptoms. *Psychological Medicine* **52**, 1458–1470 (2022) DOI:10.1017/S0033291720003244.
78. Wang, Z., Goerlich, K. S., Luo, Y., Xu, P. & Aleman, A. Social-specific impairment of negative emotion perception in alexithymia. *Soc Cogn Affect Neurosci* **17**, 387–397 (2021) DOI:10.1093/scan/nsab099.
79. Liu, C., Chai, J. W. & Yu, R. Negative incidental emotions augment fairness sensitivity. *Sci Rep* **6**, 24892 (2016) DOI:10.1038/srep24892.
80. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using {lme4}. *Journal of Statistical Software* **67**, 1–48 (2015) DOI:10.18637/jss.v067.i01.
81. Kunzetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest package: tests in linear mixed effect models. *J Stat Softw* **82**, 1–26 (2017).
82. Arel-Bundock, V., Greifer, N. & Heiss, A. How to Interpret Statistical Models Using marginaleffects for R and Python. *Journal of Statistical Software* **111**, 1–32 (2024) DOI:10.18637/jss.v111.i09.
83. Rosseel, Y. lavaan: An R package for structural equation modeling. *Journal of statistical software* **48**, 1–36 (2012).
84. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag New York, 2009).
85. Lüdecke, D. ggeffects: Tidy Data Frames of Marginal Effects from Regression Models. *Journal of Open Source Software* **3**, 772 (2018) DOI:10.21105/joss.00772.