

# Affective Abstraction Predicts Variation in Alexithymia, Depression, and Autism Spectrum Quotient

Stephanie A. Fiedler<sup>1, 2</sup>, Kent M. Lee<sup>2, 3</sup>, Erik C. Nook<sup>4</sup>, Kristen A. Lindquist<sup>5</sup>,  
Maria Gendron<sup>6</sup>, and Ajay B. Satpute<sup>2, 3</sup>

<sup>1</sup> Department of Psychology, Wayne State University

<sup>2</sup> Department of Psychology, Northeastern University

<sup>3</sup> Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital

<sup>4</sup> Department of Psychology, Princeton University

<sup>5</sup> Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill

<sup>6</sup> Department of Psychology, Yale University

Affective abstraction refers to how people conceptualize affective states in terms of category-level representations that generalize across specific situations (e.g., “fear” as evoked by heights, predators, and haunted houses). Here, we develop a novel task for assessing affective abstraction and test its relations with trait alexithymia, depression, and autism spectrum quotient. In a preregistered online study, participants completed a set of tasks in which they matched a cue image with one of two probe images based on similarity of affective experience. In a discrete emotion version of the task, the cue and target probe matched on a discrete emotion category while controlling for valence. In a valence version of the task, the cue and target probe matched on valence (i.e., pleasantness or unpleasantness). We further varied the degree of abstraction such that some judgments crossed semantic categories (e.g., a house cue with animal probes). Accuracy, as indexed by the proportion of choices that accorded with norms, predicted trait measures of alexithymia, depression, and autism quotient with medium effect sizes. We conducted an integrative data analysis by including data from three other (nonpreregistered) samples ( $N = 435$ ) and found substantial moderation by sampling population (Amazon Mechanical Turk, college students) and partial moderation by gender identity. Additional constraints on generalization include that our sample included predominantly White American adults between the ages of 23 and 64. These results provide preliminary support for the notion that affective abstraction may reflect a transdiagnostic psychological process of broad relevance to individual differences in affective processing.

**Keywords:** emotion, alexithymia, depression, abstraction, autism

**Supplemental materials:** <https://doi.org/10.1037/emo0001497.supp>

What do we mean when we say that two events both make us feel “afraid”? Consider, for example, that encountering a predator or looking down a sheer cliff may be quite distinct in terms of their sensory, cognitive, and behavioral features, and yet one might still comprehend both instances as generating an affective experience that belongs to the category “fear.” A great deal of research has focused on how people identify emotions or tell instances of

emotion apart from one another (i.e., emotion granularity or emotion differentiation; Hoemann et al., 2021; Suvak et al., 2011). However, recent theoretical advances suggest that the ability to identify or differentiate emotion may be facilitated by abstraction, or the process of categorizing distinct instances as belonging to the same category (Figure 1; Barrett, 2022; Nook et al., 2020; Satpute & Lindquist, 2019; Wilson-Mendenhall et al., 2011). In this article, we

---

This article was published Online First March 31, 2025.

C. Daryl Cameron served as action editor.

Stephanie A. Fiedler  <https://orcid.org/0000-0002-6470-7422>

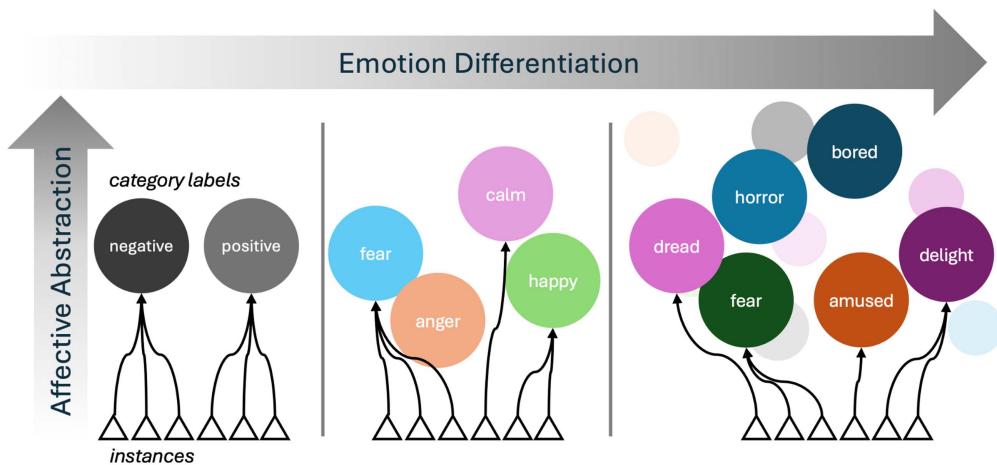
This research was supported by the Division of Behavioral and Cognitive Sciences of the National Science Foundation (Grant 2241938) awarded to Ajay B. Satpute and a National Research Service Award from the National Institute of Mental Health, National Institutes of Health (Grant F32MH122062). The authors thank Lauren Christenson and Abigail MacPherson for their contribution in the development of the study materials. They also thank Béatrice Schueler for support with data visualizations.

Stephanie A. Fiedler played a lead role in data curation, formal analysis, methodology, project administration, validation, and visualization, a supporting role in conceptualization, and an equal role in writing—original

draft. Kent M. Lee played a supporting role in conceptualization, supervision, writing—original draft, and writing—review and editing. Erik C. Nook played a supporting role in writing—review and editing. Kristen A. Lindquist played a supporting role in writing—review and editing. Maria Gendron played a supporting role in writing—review and editing. Ajay B. Satpute played a lead role in conceptualization, funding acquisition, and supervision and a supporting role in writing—original draft.

Correspondence concerning this article should be addressed to Stephanie A. Fiedler, Department of Psychology, Wayne State University, 5057 Woodward Avenue, Floor 7, Detroit, MI 48202, United States, or Ajay B. Satpute, Department of Psychology, Northeastern University, 360 Huntington Avenue, Boston, MA 02115, United States. Email: [stephanie.fiedler@wayne.edu](mailto:stephanie.fiedler@wayne.edu) or [a.satpute@northeastern.edu](mailto:a.satpute@northeastern.edu)

**Figure 1**  
*Theoretical Relationship Between Affective Abstraction and Emotion Differentiation*



**Note.** Emotion differentiation refers to how people parse their experiences (Hoemann et al., 2021). Low emotion differentiation arises when people lump their instances (depicted as triangles) into just a few emotion categories (e.g., toward the left), whereas high emotion differentiation arises when people have high granularity in how they conceptualize and categorize their emotions (e.g., toward the right). Thus, emotion differentiation refers to the specificity with which people match instances to categories. Affective abstraction refers to the process through which people develop and deploy an emotion category. Affective abstraction allows people to group together instances (depicted as triangles) in the same category despite variation in their surface features. This is a key distinction in terms as affective abstraction is involved in emotion conceptualization regardless of levels of granularity. Abstraction can be extremely broad focus (e.g., identifying similarities in valence across instances) or more narrow (e.g., identifying similarities that align with specific emotion categories). Here, we measure affective abstraction along both dimensions using the Valence-Experience Matching Task and Emotion-Experience Matching Task. We did not measure emotion differentiation in this study, although it is typically assessed using intraclass correlations of emotion ratings over time. See the online article for the color version of this figure.

focus specifically on abstraction in instances of emotion and valence. We propose that affective abstraction may reflect a transdiagnostic psychological process of broad relevance to individual differences in affective processing.

### Constructionism and Affective Abstraction

We use the term “affective abstraction” to refer to how people categorize diverse instances into a common emotion or valence category. As a theoretical construct, affective abstraction originates from a constructionist account of emotion. Constructionism proposes that emotions are enculturated as embodied concepts that are acquired during development and vary across the lifespan (Lindquist et al., 2022; Nook et al., 2017; Widen et al., 2015). Instances of emotion are thought to occur when prior experiences pertaining to emotion categories are drawn upon to “explain” ongoing sensory input (including interoceptive and exteroceptive inputs) in a predictive process (Barrett, 2017; K. M. Lee et al., 2021). The set of prior experiences that the brain uses in any given moment will naturally vary across individuals and situations. Thus, constructionist models posit that each emotion category representation is a heterogeneous collection of instances rather than a type (Barrett, 2013; Barrett & Satpute, 2019; Lindquist, Satpute, & Gendron, 2015). This notion distinguishes constructionist theory from both classical views that assume that each emotion category involves a unique and specific set of neural

and autonomic features (e.g., Ekman, 1992; Levenson, 2003; Panksepp, 2011; Saarimäki et al., 2016), and “core + variants” views that treat each emotion category as having a core type with a relatively more modest degree of variants surrounding each type (e.g., Ekman, 1992; Keltner et al., 2019).

Consistent with constructionist hypotheses, several lines of research suggest that the biological basis of emotion is highly variable and strongly depends on the person and situation (Barrett, 2017; Larsen et al., 2008; Lindquist et al., 2012; McVeigh et al., 2024; Siegel et al., 2018; Wang et al., 2024; Wilson-Mendenhall et al., 2011). These findings lead to another question, namely, what psychological mechanisms underlie how people categorize instances with heterogeneous features as belonging to the same emotion category (Barrett, 2006)? We (Satpute & Lindquist, 2019, 2021) and others (Barrett, 2017; Spunt & Adolphs, 2017; Wilson-Mendenhall et al., 2011) have proposed that abstraction serves this role. In cognitive science, abstraction captures the ability to characterize instances as similar to one another despite variation in their physical characteristics (Barsalou & Wiemer-Hastings, 2005; Borghi et al., 2017; Murphy, 2002). For example, dollar bills, coins, and electronic bank accounts, despite having a greater surface resemblance to paper, pebbles, and email accounts, respectively, are more similar to one other in an abstract sense as forms of currency. Similarly, affective abstraction refers to how instances with heterogeneous neural correlates and sensorimotor features may nonetheless share an affect or emotion category label in

common (Doyle et al., 2022; McVeigh et al., 2024; Wang et al., 2024; Wilson-Mendenhall et al., 2011).

Affective abstraction plays a distinct but complementary role to emotion differentiation or emotion granularity (Figure 1). Emotion differentiation focuses on the specificity of one's experiences. It is commonly operationalized as the mean intraclass correlation coefficient of emotion category ratings over time (Hoemann et al., 2021; Suvak et al., 2011). Lower values are interpreted as greater granularity. In comparison, affective abstraction refers to how heterogeneous instances can be grouped as similar to one another in terms of sharing an emotion or affect category in common which can occur across levels of granularity (Figure 1) and applied flexibly using more general valence categories or more specific emotion categories.

### Affective Abstraction, Alexithymia, and Constructionist Theory

Here, we hypothesize that deficits in affective abstraction underlie variation in alexithymia (Sifneos, 1996; Taylor & Bagby, 2004) or related constructs (e.g., affective agnosia and emotional awareness; Lane et al., 1990, 2015; Weissman et al., 2020) and variation in traits and symptoms of psychopathology related to alexithymia (Taylor et al., 1999). Our hypothesis is grounded in a constructionist theoretical model of emotion (Barrett, 2006, 2017; Hoemann et al., 2020; Satpute & Lindquist, 2019; Widen & Russell, 2008) in which affective abstraction is thought to be a key facet of emotion representation and understanding (Satpute & Lindquist, 2019, 2021). In theory, deficits in affective abstraction would make it difficult to link distinct instances to one another, and thus make meaning of current instances as belonging to a given category.

Alexithymia has been described as the inability to identify and describe experiences in terms of their affective or emotional qualities. We use a three-component definition of alexithymia, wherein alexithymia is considered to involve difficulty identifying feelings, difficulty describing feelings, and externally oriented thinking (Preece et al., 2017; Preece & Gross, 2023). It has previously been associated with both transdiagnostic symptoms of psychopathology (Honkalampi et al., 2001; Kim et al., 2008; Leweke et al., 2011; Taylor et al., 1999) and neurodiverse traits (e.g., autism; Kinnaird et al., 2019; Poquérusse et al., 2018). It is likely that alexithymia arises from a multitude of social, developmental, cognitive, and neural mechanisms (Booth & Happé, 2018; Díaz & Prinz, 2023; Hobson et al., 2019; Moriguchi & Komaki, 2013; Nook et al., 2015). For that same reason, any observed relationships between affective abstraction and alexithymia need not imply that said process is also of transdiagnostic relevance. Thus, we further examined subclinical trait variation in depression and autism spectrum quotient to determine whether deficits in affective abstraction are related to psychopathology and neurodiversity, respectively, even when limited to the spectrum of trait variation in a community sample.

### A Novel Experimental Task to Measure Affective Abstraction

Although impairments in the mental representation of emotion are considered to be a defining component of alexithymia (Sifneos, 1973; Taylor et al., 2016), there are limited tools for assessing

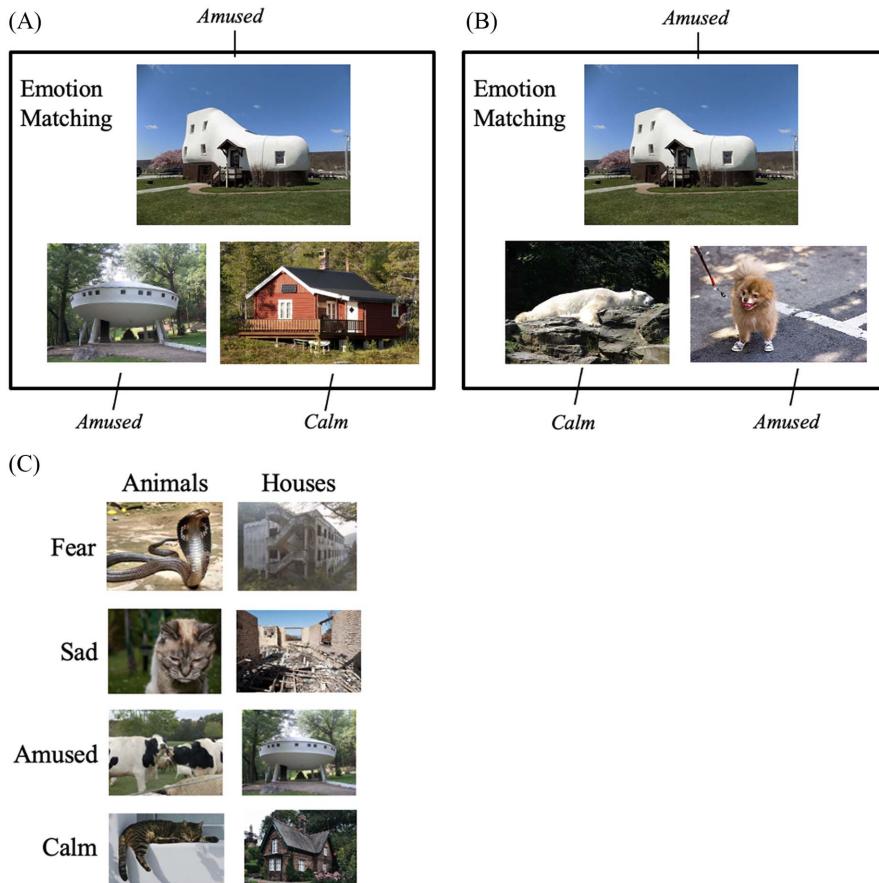
relations between affective experiences and alexithymia. This might be because of the challenge in developing performance-based benchmarks in studies focusing on affective experience (see K. M. Lee et al., 2020; Mauss & Robinson, 2009). Most studies rely on self-report data which pose interpretational challenges when it comes to alexithymia. For instance, if individuals high in alexithymia report reduced fear or sadness, it is unclear whether that is because of reduced sensitivity in affective processing or individual differences in the interpretation and use of self-report scales (for a discussion about self-report scales and emotion, see Barrett, 2004).

To address this limitation, we adapted a matching task used in semantic memory research. This task examines how people extract category-level representations from concrete instances (Bartram, 1976; for an example in social cognition, see Satpute et al., 2014). Participants are simply asked to match a cue stimulus with one of two probe stimuli, where one probe (the target) matches the cue on some abstract category dimension, and the other probe (the foil) does not share this category with the cue. By strategically manipulating the target and foil probes, the matching tasks approximate a performance-based metric of affective abstraction. Matching task paradigms have been used in emotion perception studies wherein individuals match stimuli based on emotion category (e.g., a cue face with one of two probe faces based on similarity of perceived emotion; Hariri et al., 2002; Lieberman et al., 2007; Morgan et al., 2010; Ogren & Johnson, 2020). Some of these studies reveal a small but consistent relationship between task performance (i.e., matching two faces both thought to express the same emotion) and child development (Morgan et al., 2010; Ogren & Johnson, 2020). Additional studies have also revealed a relationship between task performance and some clinically relevant traits (e.g., autism; Demenescu et al., 2010; Ulijarevic & Hamilton, 2013). Here, our aim was to examine how people match stimuli that evoke similar feelings in the self rather than the perception of emotions in others.

Participants were instructed to match a cue stimulus with one of two probe stimuli on the basis of how they felt about the stimuli. We used a normed stimulus set in which the cue and target emotion were both more strongly rated as of the same normative emotion category (e.g., fear) relative to the foil (e.g., sadness). To guard against the possibility that cues and probes could be matched based on low-level stimulus features rather than affective experience, we required participants to match cue stimuli with probes both within semantic categories (e.g., the cue and probes are all images of animals) but also across semantic categories (e.g., the cue is an image of an animal but the probes are images of houses; Figure 2A vs. Figure 2B).

We developed two versions of our task to examine affective abstraction for emotion (i.e., [fear, sad, amused, calm]) and valence (i.e., positive or negative) categories. For the Emotion-Experience Matching Task (E-EMT), the cue and target probe shared a normative emotion category (e.g., sadness), while the nontarget probe was of the same valence but different emotion category (e.g., fear; as depicted in Figure 2). For the valence version of the task, the cue and target probe were both normatively positive or negative irrespective of emotion category while the nontarget probe was of the opposite valence. The distinction between emotion and affect has recently been used to understand the nature of alexithymia such that individuals high in alexithymia may have a more pronounced deficit in emotion category differentiation rather than more global processing

**Figure 2**  
Overview of the Emotion-Experience Matching Task (E-EMT)



**Note.** An overview of the E-EMT. (A) On each trial, participants were instructed to match a cue picture (top row) to one of two probe pictures (bottom row) based on which one evoked a more similar feeling. On this sample trial, the cue matched the probes in semantic content (all are houses). (B) The cues could also involve a different semantic content. The normative emotion labels for the images in panels (A, B) are included for illustrative purposes; they were not shown to participants. Task instructions remained on the screen across all trials for each condition. (C) Sample images by content and normative emotion category conditions. Due to copyright restrictions, most images shown here are different from those used in the study. Selected images used in the tasks are demarcated by an asterisk (\*) in the attributions. The remaining images were not used in the task but selected by authors based on resemblance to those used in the study. Panel A image credit from top to bottom, left to right: (i) PLBthetoonist, CC BY-SA 4.0, via Wikimedia Commons, (ii) Joseph Novak, CC BY 2.0, via Wikimedia Commons, \* (iii) Photo by Barnabas Davoti on Unsplash. Panel B image credit from top to bottom, left to right: (i) same as Panel A; (ii) Jean-Luc 2005 at German Wikipedia, CC BY-SA 3.0, via Wikimedia Commons, via GNU Free Documentation License, Version 1.2; (iii) Victorgrigas, CC BY-SA 3.0, via Wikimedia Commons. Panel C image credit from top left to bottom right: (i) Chandan Singh from India, CC BY 2.0, via Wikimedia Commons\*, (ii) Sanandreas119, CC0, via Wikimedia Commons; (iii) Dimitri Torterat, CC BY 2.0 FR, via Wikimedia Commons; (iv) CSIRO, CC BY 3.0, via Wikimedia Commons; (v) Rikki's Refuge from Orange, Virginia, USA, CC BY 2.0, via Wikimedia Commons; (vi) same as Panel A; (vii) Image by Christophe Schindler from Pixabay\*; (viii) Photo by Abby Rurenko on Unsplash. No changes were made to Creative Commons images. See [Supplemental Section 7](#) for copyright license URLs and Unsplash and Pixabay additional attribution URLs. See the online article for the color version of this figure.

of valence (Nook et al., 2015; Preece et al., 2018). Thus, one hypothesis is that individuals high in alexithymia may show a specific deficit in abstraction, when it comes to discrete emotion categories but not when matching on valence. Alternatively, individuals high in alexithymia may exhibit a more general deficit in

affective abstraction, regardless of whether it concerns emotion or valence, in part because valence itself has also been viewed as an abstract dimension (Igaya et al., 2021; Ruba et al., 2021; Widen & Russell, 2008; also see [Figure 1](#)). In the present study, we address these possibilities by examining how error rates in emotion and

affect versions of our matching tasks relate with both alexithymia scales, in addition to examining how these relationships extend to trait variation in depressive and autistic traits, at least in neurotypical adults, to test for potential transdiagnostic implications.

### Study 1 (Preregistered)

#### Method

Table 1 provides a list of acronyms used in the sections below.

#### Sample

This study was conducted in compliance with the Northeastern University Institutional Review Board. Adult participants were recruited from Amazon Mechanical Turk (MTurk) between 2020 and 2021, and were compensated US\$8 for the completion of the study. We preregistered (AsPredicted No. 74290; [https://aspredicted.org/CT4\\_561](https://aspredicted.org/CT4_561)) a recruitment sample of  $N = 250$  prior to exclusions due to data quality control. Of participants who began the experiment, 19 participants did not complete the three tasks and one participant fell outside of the age range for inclusion and hence was not included in our targeted recruitment sample. Several pre-registered data quality control steps were taken to account for other sources of poor data quality (e.g., bots, reduced subject attention) and are detailed below. The final usable sample included 127 participants (female = 43, male = 84, nonbinary or other = 0) whose ages ranged from 23 to 64 ( $M = 36$ ,  $SD = 9.80$ ). Of those included, 82% identified as White, 10% as Black/African American, 0.8% as American Indian/Alaska Native, 7% as Asian, and 3% as multiple races; for ethnicity, 19% identified as Hispanic/Latino.

#### Stimuli

We curated a unique picture stimulus set for this experiment. Specifically, we curated a set of 80 pictures that normatively evoked four target emotions across two types of semantic contents (i.e., 10 Pictures Per Emotion  $\times$  Content Combination). The norming sample was recruited from the same population as the study sample. Target emotions included two negative valence emotions, fear and sadness, and two positive valence emotions, amusement and calmness. We selected four emotions to span the valence and arousal dimensions (fear = high arousal negative, sadness = low arousal negative, amusement = high arousal positive, calmness = low arousal positive). The picture content was either of nonhuman animals or of houses. We specifically obtained stimuli that were normed to be both high on one emotion category (e.g., sadness) and low on

another (e.g., fear; Table 2). Additional details on stimulus norming are available in the Supplemental Section 1. Notably, there was variation in the norm consistency across individual stimuli assigned to each category (see Supplemental Table S1). Thus, we also performed a norm consistency analysis. We limited analyses to trials wherein the lowest image norm exceeded different thresholds, which showed robustness of correlations between trait measures and matching accuracy on the E-EMT (see Supplemental Table S4).

#### E-EMT

Participants were instructed to match a cue picture (top row of Figure 2) to one of two probe pictures (bottom row of Figure 2) based on similarity of emotional experience. On a given trial, the cue and probe pictures were all either of positive or negative valence, but the cue and one of the probes shared the same dominant normative emotion experience. For example, an amusing cue picture was presented with an amusing probe picture and a calm probe picture (Figure 2A). There were no specific emotion words presented in the task; however, the matching instructions “Emotion Matching” appeared in the top left corner of each trial. To ensure that experience matching was not dependent on similarities in semantic or perceptual features, we manipulated the content categories of cue and probe pictures. For “same content category” trials, the cue and probe pictures were all from the same content category (e.g., all animals). For “different content category” trials, the cue picture depicted one content category (e.g., an animal), whereas the probe pictures depicted the other content category (e.g., houses).

Each of the 80 pictures was repeated three times in three different trials per task, once as a cue picture and twice as probe pictures. This procedure led to 80 potential trials for each task. However, to reduce fatigue, the stimuli were split into two lists of 40 trials, and each participant only completed 40 trials per task. Each list was balanced to contain an equal number of target emotion or valence categories for cues, an equal number of same/different content conditions for cue–probe matches, and an equal number of trials across both factors (i.e., there were five trials for each  $4 \times 2$  combination of emotion category condition with content matching condition, and 10 trials for each  $2 \times 2$  combination of valence category condition with content matching condition). Participants were randomly assigned to view stimuli from one of the two lists for each task, and the trial order was randomized for each task and participant. The task was self-paced and only moved forward when the participant clicked on the probe picture of their choice. Although the task was self-paced, participants who did not complete the task within 3 hr were automatically excluded to ensure the study was done in one sitting. There were no intertrial intervals between trials, but between tasks, there were instructions and practice trials for the subsequent task.

#### Valence-Experience Matching Task

The Valence-Experience Matching Task (V-EMT) was similar to the E-EMT except the probe pictures differed in valence from each other. For example, a positive valence cue picture (which could either be amusing or calm) would be presented with a positive valence probe picture (also either amusing or calm, counterbalanced across trials) and a negative valence probe picture (either fearful or sad). Here, there were also no specific emotion words presented in

**Table 1**  
*Reference Table for Key Acronyms*

Acronym	Phrase
E-EMT	Emotion-Experience Matching Task
V-EMT	Valence-Experience Matching Task
TAS-20	Toronto Alexithymia Scale-20 items
PAQ	Perth Alexithymia Questionnaire
CES-D	Center for Epidemiologic Studies Depression Scale
AQ	Autism quotient

**Table 2**  
Normative Ratings for Picture Stimuli

Stimulus categories	Mean rating				SD			
	Amused	Calm	Fear	Sad	Amused	Calm	Fear	Sad
Amused	.73	.11	.09	.07	.09	.07	.09	.05
Animal	.71	.10	.12	.07	.12	.07	.12	.06
House	.75	.13	.07	.06	.05	.07	.05	.05
Calm	.17	.73	.04	.05	.09	.08	.03	.05
Animal	.15	.75	.03	.07	.11	.10	.02	.05
House	.19	.72	.06	.03	.06	.06	.04	.03
Fear	.07	.07	.77	.09	.06	.06	.12	.08
Animal	.07	.05	.81	.07	.05	.04	.06	.07
House	.07	.10	.74	.10	.08	.07	.15	.09
Sad	.10	.07	.09	.74	.06	.09	.08	.16
Animal	.07	.05	.02	.85	.04	.04	.02	.05
House	.12	.09	.16	.63	.06	.12	.06	.15

Note. Ms and SDs of the final stage of norming. Values in bold emphasize correspondence between measures of normative emotion ratings with category labels. See Supplemental Section 1, for additional details on stimulus norming stages.

the task; however, the matching instructions “Positive or Negative Matching” appeared in the top left corner of each trial. Again, content categories were also manipulated within this condition such that half were same content category trials and half were different content category trials. The same lists were used to organize stimuli into trials, and the same procedures were used to randomly assign lists across participants. Participants completed 40 trials of the V-EMT.

### Content Category Matching Task

For data quality control purposes, participants also completed a simple content category matching task in which they were instructed to match a cue from one content category (e.g., house) with a probe from the same content category (e.g., another house) instead of a probe from the other category (e.g., animal). There were no specific content words (animals, houses) presented in the task; however, the matching instructions “Content Matching” appeared in the top left corner of each trial. To maintain consistency with the E-EMT and V-EMT, the semantic matching task also contained 40 trials, evenly split with animals and houses as cues. The semantic matching task was completed in between the E-EMT and V-EMT to provide conceptual separation between two tasks, both of which required focusing on affective features but along different dimensions (i.e., in terms of emotion or valence categories).

### Inventories

**Alexithymia: Toronto Alexithymia Scale and the Perth Alexithymia Questionnaire.** The Toronto Alexithymia Scale (TAS-20; Bagby et al., 1994) and Perth Alexithymia Questionnaire (PAQ; Preece et al., 2018) were used to measure trait alexithymia. Both scales measure difficulty in identifying and describing feelings and a facet for externally oriented thinking. Correspondingly, they tend to correlate highly with each other (e.g., Zahid et al., 2024). The main difference between the scales is that the PAQ includes distinct questions for negative and positive valence feelings and reconceptualizes externally oriented thinking as a tendency to avoid thinking about feelings. Participants rated the TAS-20 items (e.g., “I have feelings that I cannot quite identify,” “I do not know what is

going on inside me”) on a 1–5 Likert scale from “disagree strongly” to “agree strongly.” The range of possible scores is 20–100. Following convention, Items 4, 5, 10, 18, and 19 were reversed scored, then all responses were summed. We found a wide range of variation in scores, ranging from 25 to 81 ( $M = 55.06$ ,  $SD = 14.77$ , *interquartile range [IQR] = 44–68*). While we focus on composite TAS-20 scores, we also present findings and intercorrelations with subscale measures in Supplemental Section 3.

The PAQ has 24 items (e.g., “When I am feeling bad, I cannot make sense of those feelings,” “When I am feeling good, I cannot talk about those feelings in much depth or detail,” “Usually, I try to avoid thinking about what I am feeling”) that were rated on a 1–7 Likert scale from “strongly disagree” to “strongly agree.” The range of possible scores is 24–168. To compute the score, all responses were summed. We found a wide range of variation in scores ranging from 24 to 158 ( $M = 61.44$ ,  $SD = 24.24$ , *IQR = 48–69*). The internal consistency for both scales in our sample was high ( $\alpha = .92$ ,  $.92$ , respectively). While we focus on composite PAQ scores, we also present findings and intercorrelations with subscale measures in Supplemental Section 3.

**Center for Epidemiologic Studies Depression Scale.** The Center for Epidemiologic Studies Depression Scale (CES-D; Radloff, 1977) was used to measure trait depression. The CES-D had 20 items that were rated on a 0–3 Likert scale from “Rarely or none of the time (less than 1 day)” to “Most or all of the time (5–7 days)” (e.g., “I thought my life had been a failure,” “I felt that I could not shake off the blues even with help from my family or friends”). The internal consistency in our sample was high ( $\alpha = .93$ ). The range of possible scores is 0–60. Following conventions, Items 4, 8, 12, and 16 were reversed scored, then all responses were summed. Scores ranged from 0 to 50 ( $M = 20.72$ ,  $SD = 14.45$ , *IQR = 6–34*).

**Autism Quotient.** The autism quotient (AQ; Baron-Cohen et al., 2001) was used to measure autistic traits. The AQ has 50 items that were rated on a 1–4 Likert scale from “definitely agree” to “definitely disagree” (e.g., “Other people frequently tell me that what I have said is impolite, even though I think it is polite,” “I tend to have very strong interests, which I get upset about if I cannot pursue”). The range of possible scores is 0–50. Items, half of which

were reverse scored, were recalculated as a 1 if the participant rated any items a 3 or 4 and a 0 if participants rated any items as a 1 or 2, and then summed for a final score. The internal consistency in our sample was high ( $\alpha = .95$ ). We found a moderate range of variation in scores, ranging from 5 to 31 ( $M = 21.57$ ,  $SD = 5.80$ ,  $IQR = 19\text{--}25$ ).

### Procedure

The study was administered using Gorilla Experiment Builder (Anwyl-Irvine et al., 2020). Participants were first provided informed consent; the experiment continued for those who agreed and otherwise terminated. Participants completed questionnaires detailed in the inventories section, followed by the behavioral tasks, a demographic survey, and debriefing. Because our hypotheses primarily concerned the E-EMT, it was administered first to mitigate against habituation and experimental fatigue on this task, followed by the semantic matching task, and, finally, the V-EMT.

For each task, participants were instructed how to match the pictures (e.g., match by emotion for the E-EMT), and practice trials were included where participants were required to practice matching by the specific criteria stated. During practice trials, participants received feedback, informing them if they had selected the correct or incorrect probe on each trial, and they repeated incorrect trials. The purpose of feedback on practice trials was to ensure participants encoded the instructions. After participants successfully matched all the practice trials, they were given the option to practice the practice trials again. If they chose not to, they were asked to verify on what criteria they should match the pictures on for the current task. If participants chose the incorrect matching rule, they were taken back to the question and told to try again. If participants chose the correct matching rule, they were told they were correct. Participants were then informed of the length of the task (i.e., “40 pictures,” representing the 40 trials), and whether there were subsequent tasks. Only after these, procedures were participants shown the trials. In addition, on every trial screen, there was a description in the top left corner of what participants were supposed to match by (e.g., Emotion Matching).

### Data Quality Control

Online data collection requires rigorous data quality control. Participants who did not complete the study were excluded ( $N = 19$ ). We then followed data quality control procedures as described in our preregistration. The simple content matching task (i.e., matching house-to-house, animal-to-animal, and not house-to-animal or animal-to-house) was used as a quality check. Participants scoring less than 90% accuracy (36/40 trials) on the content matching task were excluded ( $N = 82$ ). Participants with a Cook’s  $D > 3$  when examining trait measures for consistency in typical responding by taking advantage of the expected negative association between forward and reverse coded items were excluded ( $N = 49$ ). Of the 49 participants with a Cook’s  $D > 3$  on trait measures examining forward and reverse coded items, 15 had a Cook’s  $D > 3$  on the TAS-20, 18 on the CES-D, and 16 on the AQ. The Cook’s  $D$  was determined by correlating reversed and nonreversed items on the TAS-20, CES-D, and AQ. The PAQ did not have reversed items. This procedure assessed whether participants actually answered the questions for the trait measures (e.g., if they answered a lot of low

numbers on nonreversed items then on reversed scored items, we would expect higher numbers) or just choose the same numbers for everything, effectively not actually doing the measures (e.g., both their nonreversed items and reversed items were low numbers, or both were high numbers). The nonreversed items and reversed items should be correlated so we calculated a Cook’s  $D$  to quantify any participants who were over a reasonable threshold away from the mean which would indicate they did not actually answer the trait measures and should be excluded.

Participants who scored  $\pm 3 SD$  from the mean on any of the inventories were excluded ( $N = 1$ , on the AQ). There were no participants who took unusually long to complete the task trials using the median reaction times (RT), identified using  $z \geq 3 SD$  or any univariate outliers for each measure, identified using  $z \geq 3 SD$ . There were also no participants who fell  $3 SD$  units away from the mean Mahalanobis distance when measuring multivariate outliers from multiple regressions to examine the relationship between scores. The Mahalanobis distance is a commonly used measure of distance for multivariate data. There were 18 participants who met more than one exclusion criteria. After all exclusion criteria were applied,  $N = 114$  participants were excluded leaving  $N = 136$  included. It is possible that participants did not follow task instructions and simply chose pictures on the left or on the right. This was not accounted for in our initial preregistration, and so accounting for this possibility required a slight deviation from the preregistered exclusion criteria. Nine additional participants were excluded who showed a significant left or right button press bias in response tendency, defined as  $\pm 2 SD$  from the mean picture choice. The final total of included participants was  $N = 127$ . Note that the main findings regarding correlations between error rates on the experience matching tasks and the four inventories hold when including all participants who completed the study (all  $p < .001$ ).

### Analysis Strategy

To estimate accuracy, we used agreement of a given participant’s responses with normative categories as a benchmark (e.g., on a V-EMT trial with a normatively “negative” image, participants chose the cue that is also normatively rated as “negative”). This strategy is also used in emotion perception studies and in semantic memory studies more generally insofar as meaning is a socially agreed upon phenomenon (Nook et al., 2015). To index task difficulty, we also computed the median RT for each condition and task, which provides a more robust measure of central tendency than mean RT by preventing undue influence of skewness or outliers in the distribution (Ratcliff, 1993).

To test our hypotheses, we examined Pearson’s correlations between accuracy measures on both E-EMT and V-EMT with individual difference measures (TAS-20, PAQ, CES-D, and AQ), and we conducted a principal components analysis (PCA) to examine whether intercorrelations loaded on a single dimension. Our main interests concern correlations between task variables and individual difference measures. However, since this is a novel task, we first examined task variables to see whether certain task conditions (i.e., matching across different vs. same semantic content stimuli) and measures (RT and accuracy) carried unique or redundant information for examining individual differences. We calculated four behavioral measures when considering accuracy and RT for both different/same semantic content conditions.

Because E-EMT and V-EMT tasks were, by design, acquired serially and not counterbalanced, we analyzed them independently from one another.

### Preregistered Hypotheses

We had three main preregistered hypotheses. The first concerned the effects of task conditions on performance measures when averaging across individuals. We hypothesized that the effect of content category (i.e., matching cues to same vs. different semantic content probes) on task performance measures (accuracy, RT) would depend on the task (emotion vs. valence). Our specific prediction was that matching across different content would be more demanding in the E-EMT versus V-EMT. In the preregistration, we stated we would perform a  $2 \times 2$  repeated-measures analysis of variance to test this prediction. While this analysis confirmed our prediction, for the article, we conducted and reported analyses for the tasks separately instead, since the tasks were presented serially and were not counterbalanced (by design). Second, we hypothesized a negative association between alexithymia scores with overall task performance on the E-EMT; we did not preregister predictions concerning the V-EMT. Third, we proposed that if significant correlations are observed, then we would use mediation analyses to interrelate affective abstraction performance, alexithymia, depression, and autism quotient. While these mediation analyses were significant as predicted, there are growing concerns of using mediation analysis with cross-sectional data and upon further consideration, the hypothesis could be captured by simply using PCA. Finally, we also preregistered the hypothesis that the content matching condition would have a stronger effect on performance in the E-EMT than in the V-EMT, which was confirmed using a  $2 \times 2$  repeated-measures analysis of variance with follow-up tests (for details, see [Supplemental Section 3e](#)). However, this direct comparison between tasks is limited, since we did not counterbalance the tasks over time, and so we do not include further discussion of these results here.

### Transparency and Openness

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. The pre-registered, cleaned data for all studies including Study 1 and Studies S1–S3, along with Matlab and markdown scripts for data analysis of results presented in the main article, are publicly available on Github ([Affective and Brain Sciences Lab, 2023](#)). The full stimulus materials set cannot be posted online due to copyright; a few examples

are posted for the reader in [Figure 2](#). Stimulus materials will be shared with other researchers upon request. Study 1 was preregistered ([https://aspredicted.org/CT4\\_561](https://aspredicted.org/CT4_561)).

## Results

### Task Performance Measures

For the E-EMT, we preregistered the hypothesis that matching cues to probes from different content categories would be more demanding than matching within the same content category as indexed by RT and accuracy scores. As expected, matching accuracy for the different content condition ( $M = 0.66$ ,  $SE = .01$ ) was lower than for same content condition,  $M = 0.82$ ,  $SE = .01$ ; paired samples,  $t(126) = -12.54$ ,  $p < .001$ , 95% confidence interval (CI)  $[-.19, -.14]$ , Cohen's  $d = 1.11$ . Furthermore, participants on average also took longer on different content trials ( $M = 3.31$  s,  $SE = .13$ ) than same content trials,  $M = 3.02$  s,  $SE = .11$ ; paired samples,  $t(126) = 4.45$ ,  $p < .001$ , 95% CI [.17, .43], Cohen's  $d = 0.39$ . Unexpectedly, these findings were not the same on the V-EMT. Matching accuracy for the different content condition ( $M = 0.72$ ,  $SE = .02$ ) was slightly but significantly *higher* than for same content condition,  $M = 0.68$ ,  $SE = .02$ ; paired samples,  $t(126) = 3.68$ ,  $p < .01$ , 95% CI [.02, .07], Cohen's  $d = 0.33$ . There was also no differences in median RT for matching on different content trials ( $M = 2.40$  s,  $SE = .10$ ) than same content trials,  $M = 2.42$  s,  $SE = .10$ ; paired samples,  $t(126) = -.37$ ,  $p > .05$ , 95% CI  $[-.09, .06]$ , Cohen's  $d = 0.03$ . Overall, these findings suggest that emotion abstraction may place additional demands on abstraction across contents, but valence abstraction may not.

Next, we correlated task variables to determine whether they provide unique information for examining individual differences. For the E-EMT, same-content and different-content task performance was positively correlated for both matching accuracy ( $r = .54$ ,  $p < .001$ ) and median RT ( $r = .87$ ,  $p < .001$ ). However, matching accuracy and median RT measures were not significantly correlated with one another in either the same content or different content conditions ([Table 3](#)), suggesting the absence of speed/accuracy tradeoffs. Similar findings were observed for the V-EMT. Task performance in the same content condition was highly correlated with performance in the different content condition, for both matching accuracy ( $r = .82$ ,  $p < .001$ ) and median RT ( $r = .92$ ,  $p < .001$ ; [Table 3](#)), and accuracy and median RT measures were not correlated with one another. These correlations suggest that content matching conditions may involve shared psychological processes.

**Table 3**  
E-EMT and V-EMT Correlations Between Same and Different Semantic Content Conditions

Content category	Measure	E-EMT				V-EMT			
		Same content		Different content		Same content		Different content	
		Accuracy	Median RT	Accuracy	Median RT	Accuracy	Median RT	Accuracy	Median RT
Same	Accuracy	—				—			
	Median RT	.07	—			.01	—		
Different	Accuracy	.54**	-.02	—		.82**	.08		
	Median RT	.08	.87**	-.05	—	-.06	.92**	.02	—

Note. E-EMT = Emotion-Experience Matching Task; V-EMT = Valence-Experience Matching Task; RT = reaction time.

\*\*  $p < .001$ .

### Alexithymia Correlates With Depression and AQ

Consistent with findings in prior work, there were significant relationships between alexithymia, depression, and AQ inventory scores suggestive of transdiagnostic processes shared across them. We first examined whether trait levels of alexithymia were associated with depression and AQ as expected from prior work. Pearson's correlations between trait measures showed that both TAS-20 and PAQ scores positively correlated with each other and with CES-D and AQ scores (all  $r > .37$ ,  $p < .001$ ; **Table 4**). Second, we examined whether CES-D and AQ were also correlated. Notably, CES-D and AQ were also moderately correlated ( $r = .55$ ,  $p < .001$ ).

### E-EMT and V-EMT Accuracy Correlates With Alexithymia, Depression, and AQ

Next, we tested our preregistered hypothesis that increasing alexithymia would be negatively associated with affective abstraction. Accuracy on the E-EMT negative correlated with TAS-20 ( $r = -.55$ ,  $p < .001$ ) and PAQ ( $r < -.43$ ,  $p < .001$ ), but reaction times did not (**Figure 3** and **Table 5**). In addition, E-EMT accuracy also correlated with CES-D and AQ scores (**Figure 3** and **Table 5**). These relationships held even when controlling for age and gender (all *semipartial rs*  $[-.46, -.29]$ ; all  $p < .001$ ; see **Supplemental Table S4**). They also remained when restricting accuracy scores to different content trials, same content trials, or trials for specific positive or negative emotion categories (see **Supplemental Section 3**). For the E-EMT, we also examined correlations with trait scores when only including trials with high normative consistency, which suggest that the correlations with trait measures was not driven only by trials involving stimuli with lower normative consistency (for details, see **Supplemental Table S41**).

Similar findings were observed for the V-EMT. Matching accuracy showed significant negative correlations with the TAS-20, PAQ, CES-D, and AQ (**Figure 4** and **Table 5**), even when controlling for age and gender (all  $r > -.50$ ,  $p < .001$ ; see **Supplemental Table S4**). These relationships remained when focusing only on different content trials, same content trials, or even trials for specific positive or negative emotion categories (see **Supplemental Section 3**). In both tasks, RT was not associated with trait scores (**Table 5**), suggesting that the findings could not be accounted for by a speed/accuracy trade-off.

### Principal Components Analysis

The intercorrelations between trait measures both with each other and with accuracy measures could be explained by a common

latent factor or potential distinct factors. We submitted all of these variables (along with RT scores) to a PCA. The (unrotated) PCA returned a two dimensional solution that explained 69% of the variance, one in which trait and accuracy measures loaded together and another which explained only variance in RT scores, as shown by the component matrix (**Table 6**).

## Studies S1–S3

### Overview

In addition to Study 1, we conducted three supplemental studies (Studies S1–S3) that were not preregistered. These studies had different purposes. The goal of Study S1 was to pilot test another stimulus set. Instead of using pictures of animals and houses, we assembled and normed (**Supplemental Table S2**) another stimulus set involving pictures of paintings and artistic furniture pieces. Otherwise, Study S1 was similar in recruitment sample (MTurk) and study procedures. The goals of Studies S2 and S3 were twofold: to test for behavioral main effects of the task to plan for a future functional magnetic resonance imaging study and to inform performance benchmarks for the semantic matching task (e.g., to help exclude bots). For the latter reason in particular, Studies S2 and S3 were recruited from a different sampling population, specifically, undergraduate students at Northeastern University. Study S2 used the animals/houses stimulus set, and Study S3 used the furniture/paintings stimulus set.

The sampling population for Studies S2 and S3 (college students) was more homogeneous and lower in trait scores than in Studies 1 and S1 (**Supplemental Table S3**; cf. **Supplemental Sections 4b–i** and **5a–i**). It is possible that the relationship between affective abstraction and trait measures is robust to the sampling population. However, students from Northeastern University are also likely to differ from the online sampling population and the general U.S. population on variables that are related to trait variables (e.g., socioeconomic status, education; Honkalampi et al., 1999; also, intelligence quotient; Bagby et al., 1986), which may mitigate the relationships between trait variables and social cognitive functions, too. For these reasons, we did not make *a priori* hypotheses that the college student sample would replicate the correlations between task performances with trait measures. The findings from these studies may be informative for establishing boundary conditions for generalization across sampling populations (MTurk, college students), stimulus sets (animals/houses, paintings/furniture), and gender identity. Here, we present an integrative data analysis involving all four data sets (Study 1 and Studies S1–S3) to provide a statistical summary of the findings and moderator effects by stimulus set (animals/houses, furniture/paintings), sampling population (MTurk, college students), and gender identity. Specific analyses at the level of each study are presented in **Supplemental Materials**.

### Method

#### Samples, Stimuli, and Procedures

We aimed to recruit at least 100 participants prior to exclusions per study based on convenience sampling. Sample sizes were not based on *a priori* power calculations, and the studies were not preregistered. Study S1 recruited MTurk participants ( $N = 80$  participants after exclusions; female = 29, male = 50, nonbinary or

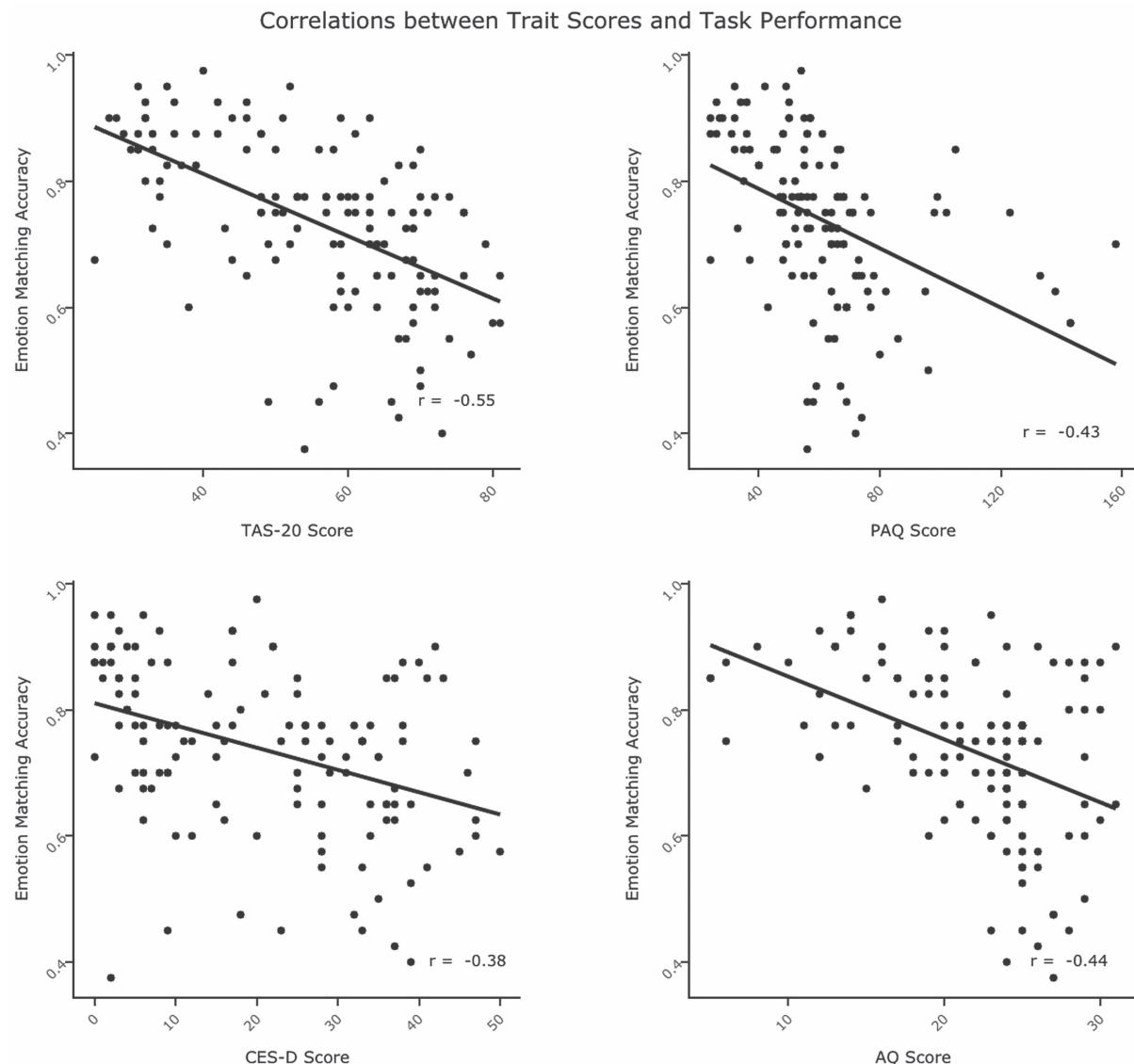
**Table 4**

Correlations Between Trait Measures

Trait measure	1	2	3	4
1. Alexithymia (TAS-20)	—			
2. Alexithymia (PAQ)	.76**	—		
3. Depression (CES-D)	.71**	.64**	—	
4. Autism quotient (AQ)	.54**	.37**	.55**	—

Note. TAS-20 = Toronto Alexithymia Scale 20 items; PAQ = Perth Alexithymia Questionnaire; CES-D = Center for Epidemiologic Studies Depression Scale.

\*\*  $p < .001$ .

**Figure 3***Study 1: Correlations Between Trait Scores and Emotion Task Performance*

**Note.** Emotion-Experience Matching Task accuracy performance is negatively correlated with increased alexithymia (TAS-20; PAQ), depression (CES-D), and autism quotient (AQ). TAS-20 = Toronto Alexithymia Scale 20 items; PAQ = Perth Alexithymia Questionnaire; CES-D = Center for Epidemiologic Studies Depression Scale.

other = 0, no response = 1; age:  $M = 35$ ,  $SD = 8.46$ , range [24, 62]) completed the same set of tasks as in Study 1 with this other stimulus set. Study S2 ( $N = 121$  participants after exclusions; female = 99, male = 20, nonbinary or other = 2; age:  $M = 18.37$ ,  $SD = .73$ , range [18, 22]) was presented with the animals/houses stimulus set. Study S3 ( $N = 107$  participants after exclusions; female = 91, male = 15, nonbinary or other = 1; age:  $M = 18.56$ ,  $SD = 1.4$ , range [18, 27]) was presented with the furniture/paintings stimulus set. Data quality control and exclusion criteria are described in the [Supplemental Materials](#).

### Analysis Strategy

We conducted a combined analysis using data across all four studies in a series of integrative data analyses (Curran & Hussong, 2009). We treated sample and stimulus sets as fixed properties at the level of each individual by including these as fixed effects in the integrative data analyses. Both variables, sampling population and stimulus set, were treated as dummy-coded variables such that one level of each variable served as a reference level (online sample and animals/houses stimulus set). Across the eight integrative data

**Table 5**  
*Correlations Between Accuracy and Trait Measures*

Trait measure	E-EMT		V-EMT	
	Accuracy	Median RT	Accuracy	Median RT
Alexithymia (TAS-20)	-.55**	.12	-.62**	.09
Alexithymia (PAQ)	-.43**	.14	-.58**	.13
Depression (CES-D)	-.38**	.05	-.46**	.01
Autism quotient (AQ)	-.44**	-.06	-.41**	-.04

Note. E-EMT = Emotion-Experience Matching Task; V-EMT = Valence-Experience Matching Task; RT = reaction time; TAS-20 = Toronto Alexithymia Scale 20 items; PAQ = Perth Alexithymia Questionnaire; CES-D = Center for Epidemiologic Studies Depression Scale.

\*\* $p < .001$ .

analysis models, E-EMT or V-EMT accuracy was treated as the outcome, and one of the four individual difference measures (TAS-20, PAQ, CES-D, AQ) was treated as an additional fixed-effect predictor, with interactions included for both sample and stimulus set. We included the self-reported gender identity of our participants as a fixed effect that interacted with the individual difference variables to predict either E-EMT or V-EMT. We only included women and men in this analysis, since there were too few individuals in other gender identity categories. To simplify the models, we did not include interactions between gender with sample or stimulus set, because we were principally interested in whether gender moderated the relationships between our individual difference variables and task performance.

## Results

A sensitivity analysis assuming  $\alpha < .05$  and power  $(1 - \beta) > .80$  in multiple regression analyses with eight tested predictors (main effects and interactions as detailed above) indicated that the total sample size ( $N = 435$ ) was powered to detect small–medium size effects (Cohen's  $f^2 \geq .035$ ;  $\beta \geq .184$ ). Table 7 presents simple correlations between individual difference measures and matching accuracy across all studies. Combining data across studies, the results of the integrative data analysis, which are summarized in Table 8, showed that all main individual difference measures (TAS-20, PAQ, CES-D, and AQ) negatively correlated with matching accuracy on both V-EMT and E-EMT (all  $ps < .001$ ). These findings were observed in both men and women, albeit the slope was slightly lower for women ( $\beta$ s ranged from  $-.19$  to  $-.41$ ,  $ps < .05$ ) than men ( $\beta$ s ranged from  $-.47$  to  $-.68$ ,  $ps < .001$ ; see Supplemental Figures S3, S5, S7, and S11).

These effects were also moderated by sampling population in both the V-EMT and E-EMT. The correlation between those individual difference measures and matching accuracy in the V-EMT was significant in the MTurk samples ( $\beta$ s ranged from  $-.19$  to  $-.46$ ,  $ps < .05$ ), but not the student samples ( $\beta$ s ranged from  $.05$  to  $-.14$ ,  $ps > .05$ ). The correlation between those individual difference measures and matching accuracy in the E-EMT was significant in the MTurk samples ( $\beta$ s ranged from  $-.28$  to  $-.44$ ,  $ps \leq .001$ ), but not the student samples ( $\beta$ s ranged from  $.02$  to  $-.06$ ,  $ps > .05$ ), for TAS-20 and AQ measures; there was no moderation for CES-D

and PAQ measures. We report the full results of all integrative data analysis regression models in Supplemental Tables S33–S40 (see Supplemental Section 8) and plot all interactions in Supplemental Figures S1–S11 (see Supplemental Section 9).

## Discussion

This study asked whether affective abstraction—the process of relating distinct instances to one another based on the similarity of their feelings—would predict trait variation in alexithymia, depression, and the autism quotient. We examined affective abstraction using a novel task in which participants were required to match experiences evoked by heterogeneous stimuli on the basis of their affective similarity. Our findings show that accuracy on affective experience matching tasks correlate with all three traits. In Study 1, these findings were robustly observed across task conditions, including within and across semantic content conditions, and across both emotion and valence categories. A PCA revealed that accuracy and trait scores loaded on a single dimension.

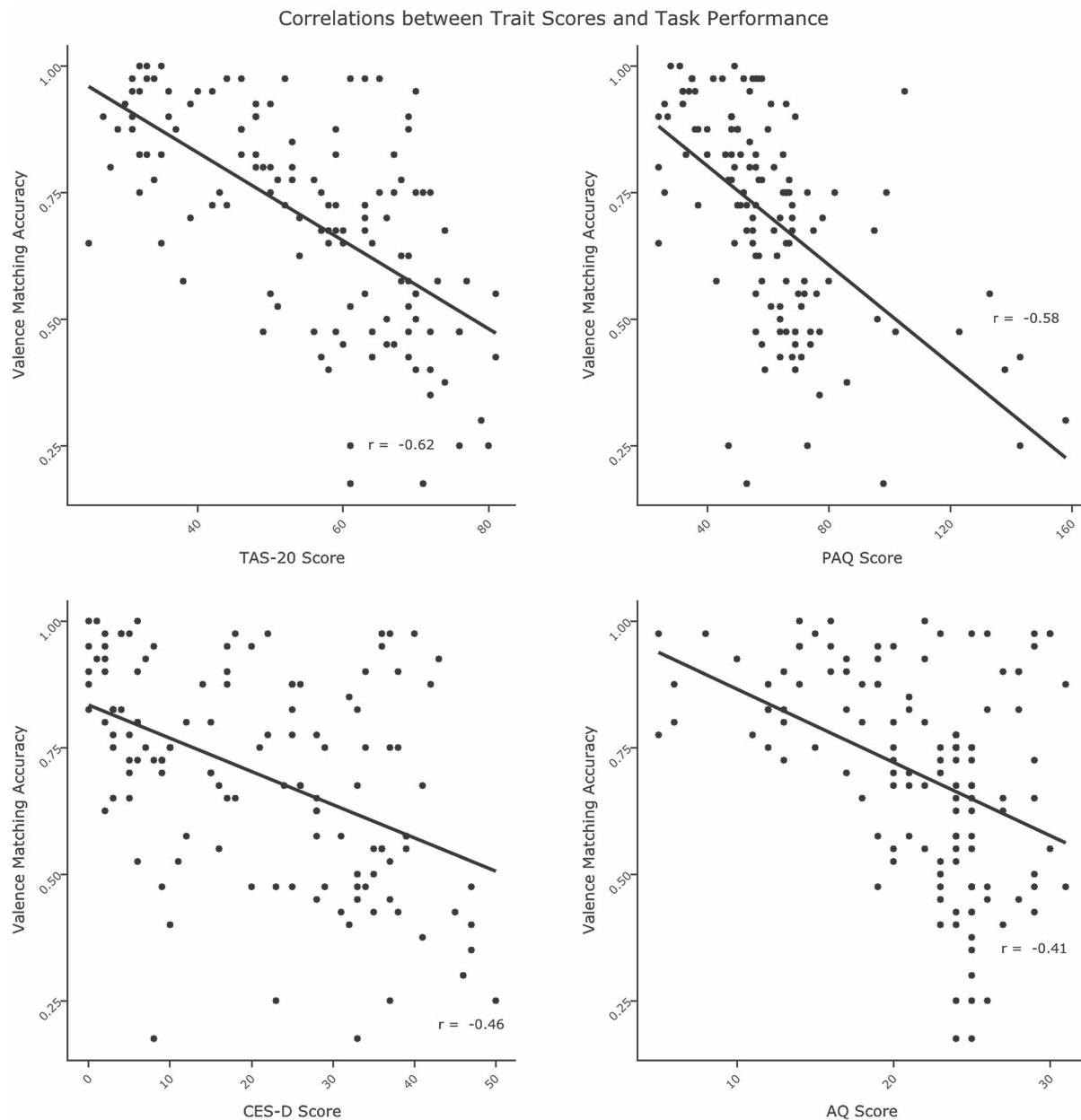
We also conducted an integrative data analysis that included three nonpreregistered studies (along with Study 1) that used different sampling populations and stimulus sets. The integrative data analysis showed that the relationships between accuracy scores and trait measures were significant overall, but also strongly moderated by sampling population. Relationships were robustly observed in the MTurk samples but not the college student samples. Notably, the distribution of trait measures included higher scores in the MTurk samples than in the college student samples. Gender identity also partially moderated the effects; relationships were stronger in men than women.

The integrative data analysis also suggests that these main findings generalize across different stimulus sets including the animals/houses set used in Study 1 and the furniture/paintings set used in Study S1. Taken together, the findings of preregistered Study 1 and the integrative data analysis suggest that affective abstraction may serve as a transdiagnostic psychological process of broad relevance to trait alexithymia and also clinical conditions (i.e., depression) and neurodiverse traits (i.e., autism) associated with alexithymia, but also that this relationship is moderated by sampling population (i.e., whether the sample includes high values on trait scores and/or involves a college student population) and gender identity.

While the integrative data analysis focused on how the main hypotheses were moderated by the characteristics of the study samples, there were also some notable differences between Studies 1 and S1, when parsing the trials into finer categories. In particular, Study 1 showed that correlations between trait scores and matching accuracy were relatively robust to the emotion category of the cue stimulus, whereas Study S1 appeared to show some variation such that the correlations varied by emotion cue category and task. We point out and discuss these observations in Supplemental Section 4d–i.

## Affective Abstraction and Alexithymia

Alexithymia was initially developed to identify individuals who, owing to difficulty in the ability to identify and describe their feelings, may not benefit from psychotherapy (Sifneos, 1973). Research on alexithymia rapidly accumulated over the past several decades showing that alexithymia is robustly associated

**Figure 4***Study 1: Correlations Between Trait Scores and Valence Task Performance*

*Note.* Valence-Experience Matching Task accuracy performance is negatively correlated with increased alexithymia (TAS-20; PAQ), depression (CES-D), and autism quotient (AQ). TAS-20 = Toronto Alexithymia Scale 20 items; PAQ = Perth Alexithymia Questionnaire; CES-D = Center for Epidemiologic Studies Depression Scale.

with a broad variety of clinical conditions and neurodiverse traits (Aaron et al., 2019; Habibi Asgarabad et al., 2023; Hemming et al., 2019; Honkalampi et al., 2022; Kinnaird et al., 2019; Sifneos, 1996; Taylor et al., 1999). However, less work has examined the psychological mechanisms contributing to alexithymia (Hegefledt et al., 2023; Hobson et al., 2019). Our hypothesis, inspired by constructionist models of emotion, suggests that impairments in affective abstraction would make it difficult to construe ongoing sensory experience in

relation to prior knowledge (i.e., similar instances in the past). Individuals who have difficulty in affective abstraction may fail to reap the benefits of using prior knowledge to facilitate sensory processing and guide behavior in future instances. Here, our findings provide the first set of empirical findings in line with this hypothesis.

The results showed that emotion and valence EMT accuracies correlated with two different scales for alexithymia, the TAS-20 and the PAQ, suggesting robustness to the alexithymia construct. The

**Table 6**  
*Component Matrix From Principal Component Analysis*

Measure	Component	
	1	2
Alexithymia (TAS-20)	<b>0.90</b>	0.05
Alexithymia (PAQ)	<b>0.82</b>	0.12
Depression (CES-D)	<b>0.80</b>	-0.03
Autism quotient (AQ)	<b>0.68</b>	-0.17
E-EMT accuracy	<b>-0.71</b>	0.13
V-EMT accuracy	<b>-0.78</b>	0.09
E-EMT median RT	0.10	<b>0.92</b>
V-EMT median RT	0.08	<b>0.92</b>

*Note.* Loadings >.6 have been presented in bold. TAS-20 = Toronto Alexithymia Scale 20 items; PAQ = Perth Alexithymia Questionnaire; CES-D = Center for Epidemiologic Studies Depression Scale; E-EMT = Emotion-Experience Matching Task; V-EMT = Valence-Experience Matching Task; RT = reaction time.

TAS-20 is more closely aligned with the psychoanalytic traditions from which the construct of alexithymia was invented (Sifneos, 1973), wherein people high in alexithymia not only have difficulty in identifying and describing feelings but also exhibit a tendency toward external oriented thinking (Bagby et al., 1994; Schroevers et al., 2022). This last dimension is sometimes overlooked in recent appraisal or cognitivist perspectives. However, it does capture a preference for concrete over abstract content, as indicated by the items (e.g., “I prefer talking to people about their daily activities rather than their feelings”; “I prefer to watch ‘light’ entertainment shows rather than psychological dramas”). Findings from Study 1 were robust to these TAS-20 subscales (Supplemental Section 3c). The construction of the PAQ was more specifically influenced by appraisal views of emotion and included separate scales for alexithymia based on valence (i.e., difficulty in identifying and describing negative feelings or positive feelings), and items that measure a preference to avoid focusing on internal feelings (Preece et al., 2017). The findings were again robust to these subscales of the PAQ (Supplemental Section 3d).

Our findings are unlikely to be accounted for by task difficulty. In Study 1, reaction time was not associated with accuracy measures or trait scores. The PCA showed that reaction time loaded on its own dimension that was separate from variance related to accuracy and

trait scores. Certain task conditions seemed to be more difficult than others in terms of reaction time, such as the E-EMT relative to the V-EMT, or making matches across versus within semantic content conditions in the E-EMT. But these potential differences in difficulty did not appear to moderate any of the correlates with trait scores.

Deficits in affective abstraction may be due to deficits in emotion identification in response to individual stimuli (i.e., making ratings consistent with cultural norms). Extant research has not yet demonstrated a robust association between emotion identification and alexithymia, however. A review of emotion perception studies found that often fewer than 30% of studies showed significant relationships between alexithymia and emotion identification of faces for a particular normative category (e.g., fear, anger, sadness); the vast majority of comparisons showed no significant relationship with alexithymia and emotion identification of faces (Grynbarg et al., 2012). We also examined behavioral findings from another review which focused on neuroimaging studies, but in which many studies used extreme groups designs to maximize power (as summarized by van der Velde et al., 2013). Again, there was no consistent evidence that alexithymia was associated with reaction time or differences in emotion judgments, and the vast majority of behavioral comparisons showed no significant relationship between alexithymia and affective ratings. The variable effects in the literature suggest that there are likely to be moderating factors that inform the relationship between emotion identification and alexithymia including the complexity or performance demands of the emotion identification task (e.g., Cook et al., 2013; Lane et al., 2000; Nook et al., 2015; Parker et al., 2005; Prkachin et al., 2009).

We did not include a separate emotion identification task due to concerns with fatigue. However, insofar as “emotion identification” is essentially agreement with cultural norms (Gendron, 2017), it is possible that alexithymia was associated with greater deviations in experiences for individual stimuli than the norm, which would subsequently contribute to differential matching effects. To address this issue, we conducted an additional analysis for Study 1, where we excluded trials with lower normative consistency (see Supplemental Table S41). Correlations between alexithymia and E-EMT accuracy remained robust even for stimuli that had high levels of normative agreement. While this analysis cannot rule out the possibility that differential normative agreement may

**Table 7**  
*Correlations Between Individual Difference Measures and Matching Accuracy Across All Studies*

Individual difference measure	E-EMT accuracy				V-EMT accuracy			
	Study 1 <sup>†</sup>	Study S1	Study S2	Study S3	Study 1	Study S1	Study S2	Study S3
Alexithymia (TAS-20)	-.55***	-.40***	-.08	-.10	-.63***	-.47***	-.10	-.03
Alexithymia (PAQ)	-.43***	-.23*	-.18*	.02	-.58***	-.29**	-.24**	-.10
Depression (CES-D)	-.38***	-.43***	-.05	-.02	-.46***	-.37***	-.02	-.02
Autism quotient (AQ)	-.44***	-.25*	-.02	-.001	-.41***	-.42***	-.03	-.03

*Note.* *Studies:* Study 1 = <sup>†</sup>preregistered; Recruitment: Amazon Mechanical Turk; Stimuli: animals/houses; Recruitment: Amazon Mechanical Turk; Stimuli: furniture/paintings; Recruitment: Northeastern Sona Systems, college students; Stimuli: animals/houses; Recruitment: Northeastern Sona Systems, college students; Stimuli: furniture/paintings; *Tasks:* E-EMT = Emotion-Experience Matching Task; V-EMT = Valence-Experience Matching Task. *Measures:* TAS-20 = Toronto Alexithymia Scale 20 items; PAQ = Perth Alexithymia Questionnaire; CES-D = Center for Epidemiologic Studies Depression Scale. Additional details for Studies S1–S3 are presented in the *Supplemental Materials*.

\**p* < .05. \*\**p* < .01. \*\*\**p* < .001.

**Table 8**

*Summary of Integrative Data Analysis Results Predicting Matching Accuracy on Task From Each Individual Difference Measures, While Controlling for Sample Type, Gender, and Stimulus Set and Their Interaction With the Individual Difference Measure*

Outcome	Individual difference predictor	Moderation effect
V-EMT	CES-D	Sample (student samples n.s.), gender (women smaller slope)
	PAQ	Sample (student samples n.s.), stimulus set (furniture/paintings had a smaller slope)
	TAS-20	Sample (student samples n.s.), gender (women smaller slope)
	AQ	Sample (student samples n.s.), gender (women smaller slope)
E-EMT	CES-D	None
	PAQ	None
	TAS-20	Sample (student samples n.s.)
	AQ	Sample (student samples n.s.), gender (women smaller slope)

Note. V-EMT = Valence-Experience Matching Task; CES-D = Center for Epidemiologic Studies Depression Scale; n.s. = effects that do not provide credible evidence of a statistically significant effect; PAQ = Perth Alexithymia Questionnaire; TAS-20 = Toronto Alexithymia Scale 20 items; AQ = autism quotient; E-EMT = Emotion-Experience Matching Task.

contribute to the findings, they do suggest that it may not be the only factor.

Even if deficits in emotion identification, or deviations in normative ratings, could account for the findings with the E-EMT, it is unlikely that this would explain the correlation between alexithymia with the V-EMT. Several prior studies have already shown that alexithymia is not associated with affect identification judgments pertaining to valence (Berthoz et al., 2002; Bird et al., 2010; Eichmann et al., 2008; Kano et al., 2003, 2007; Kugel et al., 2008; B.-T. Lee et al., 2011; Mantani et al., 2005; Mériau et al., 2006). Collectively, the lack of robust associations between alexithymia and emotion or affect identification impairments, at least as measured using typical laboratory task paradigms, makes it unlikely that these variables would account for the findings in our study.

Notably, most studies on emotion identification, including those referred to above, focus on emotion perception rather than emotion experience. Theory and certain lines of research suggest that construction of one's own and others' emotions may share overlapping processes (Erbas et al., 2016; Israelashvili et al., 2019; Skerry & Saxe, 2014). Nonetheless, it would be ideal to test the full set of relationships between emotion identification and abstraction in both perception and experience. Here, our work may be combined with prior work in the emotion perception literature (Lane et al., 1990; Morgan et al., 2010; cf. Skerry & Saxe, 2014) to develop parallel task paradigms in future work.

### ***Affective Abstraction as a Transdiagnostic Psychological Process***

A central goal of our study was to test whether affective abstraction may serve as a transdiagnostic psychological process for understanding trait variation in affect dysregulation. Our results provide preliminary support with this view. In addition to alexithymia, affective abstraction was associated with variation in both psychopathology (depression) and neurodiversity (AQ). Moreover, the PCA showed that a single dimension explained individual differences in accuracy on abstraction tasks, alexithymia, depression, and AQ. Reaction time loaded on its own on a second dimension suggesting that these associations are not merely due to task difficulty. Our study is limited; in that, we did not focus on individuals who have received clinical diagnoses. However, variation in depression and autism is now more commonly thought of as

a spectrum. While our findings are consistent with a transdiagnostic mechanism, it would be important to extend these findings in future studies drawing on samples from clinical populations.

Our study was inspired by constructionist theory (Barrett, 2013, 2017; Barrett & Satpute, 2019; K. M. Lee et al., 2021; Lindquist et al., 2022; Lindquist, MacCormack, & Shablack, 2015; Nook et al., 2017; Widen et al., 2015), which provides a mechanistic rationale regarding why affective abstraction may play a role in psychopathology and neurodiversity. Constructionist theory builds upon the view that the brain is organized to make predictions of its sensory inputs (i.e., from the external world and from the body) using prior experiences (as acquired through learning and memory; Allen & Friston, 2018; Clark, 2013; Hutchinson & Barrett, 2019). From this account, instances of emotion (or other categories of mental experiences) are psychological constructions that occur when prior instances pertaining to an emotion category are drawn upon to predict an incoming constellation of sensory input (Barrett, 2017; Barrett & Simmons, 2015; K. M. Lee et al., 2021; Seth, 2013; for related accounts, see Smith, Lane, et al., 2019; Smith, Parr, & Friston, 2019). Insofar as prior instances are not identical to the present sensory moment, emotion construction involves some degree of affective abstraction (Barrett, 2017; Satpute & Lindquist, 2019).

Based on this account, we suggest that variation in affective abstraction may take two forms: a tendency to draw on prior instances that are only very concretely related to the situation at hand (i.e., undergeneralization or concretization; Nook et al., 2021) or the opposite tendency of drawing on a surplus of only tangentially related prior instances (i.e., over generalization; e.g., see Dunsmoor & Paz, 2015). Either case would lead to suboptimal use of prior experiences in meaning making of ongoing sensory processing. Consistent with this notion, children tend to define emotion concepts by referring to specific instances in comparison to young adults, who use more general or abstract definitions that could refer to many collections of instances (Nook et al., 2020; Widen & Russell, 2008). While abstraction here was based on normative performance, which we assume reflects a middle ground between over and undergeneralization, a variant of this task paradigm could be used to examine different tendencies in how people construe their affective experiences.

One question that is not addressed in this study is whether affective abstraction is distinct from abstraction in other domains.

We did not include a nonaffective abstraction condition, because we have no theoretical reason to believe that affective abstraction is a unique cognitive process that is distinct from “nonemotion” conceptual abstraction. Rather, we view abstraction as a general process with variation emerging based on domain expertise. Indeed, prior work examining abstraction in other domains points to certain brain regions that are shared when engaging in more abstract processing irrespective of the content (namely, portions of the “default mode” network; [Satpute & Lindquist, 2019](#)). In that sense, we view affective abstraction as just one domain of abstraction, but perhaps a particularly important domain of expertise for socialization ([Spunt & Adolphs, 2019](#)). Future work may examine whether and how trait variation relates with affective abstraction or abstraction across psychological content domains.

### **Affective Abstraction and Emotion Expertise**

As a novel theoretical construct, it is important to position affective abstraction in relation to previously explored constructs in affective science related to “emotion expertise” more broadly. In a recent scoping review of articles published since 1927, [Hoemann et al. \(2021\)](#) identified 15 constructs that fall under the umbrella term of emotion expertise, which they define as individual differences in a “range of competencies related to understanding and experiencing emotions” (p. 1159) that may contribute to well-being. Included in their review were articles that examined alexithymia (e.g., [Sifneos, 1973](#)), emotional awareness and clarity (e.g., [Boden & Thompson, 2017](#); [Lane et al., 1990](#); [Thompson et al., 2009](#)), emotional intelligence (e.g., [Mayer et al., 2003](#)), emotion granularity (e.g., [Smidt & Suvak, 2015](#)), and more. A conceptual and empirical analysis of these constructs showed that they were interrelated with each other but also irreducible to a single construct. Hoemann et al. proposed that research on emotion expertise will benefit from an emergent variable model, wherein emotion expertise is not reducible to a singular latent cause, but rather that each construct (or the multiple latent dimensions characterizing each construct) enlightens a particular aspect of it (also see, [Coan, 2010](#)).

Notably, affective abstraction is conceptually distinct from these other constructs, and the method of measurement differs in important ways. Most of these constructs are measured by analyzing variation in emotion ratings or emotion word usage over time (for reviews, see [Dejonckheere et al., 2019](#); [Hamaker et al., 2015](#); [Houben et al., 2015](#); [Kuppens & Verduyn, 2017](#); [Lindquist & Barrett, 2008](#)). For instance, emodiversity quantifies how many different emotion words are endorsed over a given window of time ([Quoidbach et al., 2014](#)), emotion fluency quantifies how many emotion words a person can retrieve in a minute ([Hegefeld et al., 2023](#)), and emotion granularity quantifies how correlated emotion ratings are over time ([Tugade et al., 2004](#)). In contrast, we quantified how well people match stimuli in terms of their emotional similarity (rather than analyzing variation in emotion ratings). Theoretically speaking, whereas most previous constructs are focused on the structure of emotion representations (or frequency of emotion occurrences), affective abstraction refers to processes involved in identifying emergent affective similarities despite surface-level variation. It would be of interest in future works to examine relationships between affective abstraction and other constructs pertaining to emotion expertise to see if they predict shared or unique variance in well-being.

### **Constraints on Generalization**

**Sampling Population.** The integrative data analysis showed that the observed correlations were stronger in online MTurk samples than in undergraduate students at Northeastern University. Online samples can be more diverse on demographic and individual difference dimensions than college student samples ([Huff & Tingley, 2015](#)). The distribution of scores on trait measures of alexithymia, depression, and autism was much higher in the MTurk samples than the college samples (see [Supplemental Materials](#) for details). While our study did not specifically target clinically diagnosed participants, several participants for the MTurk samples fell into the high range on these self-reported trait scores. Another factor to consider is gender. The two MTurk studies included ~65% men/35% women, whereas the two college student samples included ~84% women/16% men. While the main findings were significant across both genders, they were also stronger in men than women (see [Supplemental Materials](#) for details), suggesting that gender may at least partially account for differences in findings between MTurk and college student populations. Finally, MTurk and college students are both particular sampling populations. Future work examining representative or stratified samples of the U.S. population and broader global population is needed to further generalize these findings.

**Design Characteristics of the Affective Experience Matching Task.** The properties of the experience matching tasks also warrant discussion. First, to ensure participants were not simply relying on low-level visual features to make their judgments, we were able to systematically vary whether the cue was of the same or different, semantic content from the probes. Second, matching tasks do not explicitly require the use of language or self-report scales. While self-report formats offer more flexibility in capturing subjective emotional experience, they also pose their own difficulties especially when examining translational or developmental populations in which language may not be as well-developed. Here, matching tasks can balance these constraints since the cue and target probe need not match perfectly, but need only be more related to one another on a dimension of interest (i.e., affective experience) than the nontarget probe. This quality of the matching task captures a quintessential aspect of affective abstraction, which is to group together instances while overlooking their situational heterogeneities. While we examined affective abstraction using an explicit matching task, this task could readily be adapted to use indirect measures (e.g., repetition priming or repetition suppression tasks, or eye-gaze paradigms).

**Alexithymia Measures.** A potential limitation of the TAS-20 is that it includes items that also reflect psychological or psychosomatic symptoms, leading some researchers to question the extent to which it is a redundant measure of mental health concerns ([Leising et al., 2009](#); [Marchesi et al., 2014](#); [Preece et al., 2020](#)). However, these limitations are perhaps offset by inclusion of the PAQ, which was explicitly designed to address these limitations of the TAS-20 (e.g., by measuring difficulties in processing positive emotions, too).

**Causal Mechanisms.** We used a cross-sectional, correlational design to test for hypothesized relationships between affective abstraction with individual difference measures. While this study design is useful for addressing our main hypotheses, it does not reveal the temporal unfolding of these relationships. In particular, future work using longitudinal study designs may examine whether

deficits in affective abstraction precede, antecede, or run concurrently with alexithymia, and relatedly, with depressive or autistic traits.

Future work may also explore the relationship between affective abstraction, as measured by our experience matching tasks, and other psychological processes that might facilitate affective abstraction. For instance, it is plausible that affective abstraction relies on working memory for affective information (Mikels & Reuter-Lorenz, 2019; Schweizer et al., 2013; Waugh et al., 2014) or on the host of processes that are thought to underlie the ability to attend to and categorize feelings (Nook et al., 2021; Satpute et al., 2013; Smith et al., 2018). Affective abstraction may also involve having fluent semantic associations between affective or emotion concepts (e.g., “fear”) with image content (e.g., content involving looming heights, predators, or public speaking). Images activate learned associations between image content and semantic representations of affect (e.g., *knowing* that something typically elicits displeasure without deeply *feeling* displeasure in the moment; Itkes et al., 2017; Itkes & Kron, 2019). This distinction may be important for understanding how people group together diverse instances of affect and emotion despite their concrete or surface, sensory motor differences (Barrett, 2006, 2017; Lindquist, MacCormack, & Shablock, 2015; Satpute & Lindquist, 2019).

## Conclusion

In sum, psychological representations of affect and emotion involve both emotion differentiation and instance integration through abstraction. While much research has examined the former, far less has examined the latter. Here, we proposed that instance integration requires abstracting across certain situation dependent features while representing the affective or emotional qualities that they share. Drawing on constructionist theory, we hypothesized and found evidence consistent with the notion that affective abstraction would be important for how people construct emotional experiences, such that deficiencies in affective abstraction would contribute to affect dysregulation as occurs in alexithymia, depression, and AQ traits. These findings provide preliminary evidence that affective abstraction may serve as a transdiagnostic construct of relevance to mental health conditions and neurodiverse populations. Our findings also show that the sampling population (online MTurk vs. college student samples) and gender identification moderated the relationships between affective abstraction and individual differences in alexithymia, depressive symptomatology, and autistic characteristics. These results suggest that future work should pay greater attention to the role of sample characteristics in the generalizability of results across studies.

## References

- Aaron, R. V., Fisher, E. A., de la Vega, R., Lumley, M. A., & Palermo, T. M. (2019). Alexithymia in individuals with chronic pain and its relation to pain intensity, physical interference, depression, and anxiety: A systematic review and meta-analysis. *Pain*, 160(5), 994–1006. <https://doi.org/10.1097/j.pain.0000000000001487>
- Affective and Brain Sciences Lab. (2023). *Emotion\_ExperienceMatchingTask* [Data set and computer software]. [https://github.com/ABS-Lab/Emotion\\_ExperienceMatchingTask](https://github.com/ABS-Lab/Emotion_ExperienceMatchingTask)
- Allen, M., & Friston, K. J. (2018). From cognitivism to autopoiesis: Towards a computational framework for the embodied mind. *Synthese*, 195(6), 2459–2482. <https://doi.org/10.1007/s11229-016-1288-5>
- Anwyl-Irvine, A. L., Massonié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, 52(1), 388–407. <https://doi.org/10.3758/s13428-019-01237-x>
- Bagby, R. M., Parker, J. D., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia Scale—I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, 38(1), 23–32. [https://doi.org/10.1016/0022-3999\(94\)90005-1](https://doi.org/10.1016/0022-3999(94)90005-1)
- Bagby, R. M., Taylor, G. J., & Ryan, D. (1986). Toronto Alexithymia Scale: Relationship with personality and psychopathology measures. *Psychotherapy and Psychosomatics*, 45(4), 207–215. <https://doi.org/10.1159/000287950>
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, malesand females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, 31(1), 5–17. <https://doi.org/10.1023/a:1005653411471>
- Barrett, L. F. (2004). Feelings or words? Understanding the content in self-report ratings of experienced emotion. *Journal of Personality and Social Psychology*, 87(2), 266–281. <https://doi.org/10.1037/0022-3514.87.2.266>
- Barrett, L. F. (2006). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review*, 10(1), 20–46. [https://doi.org/10.1207/s15327957pspr1001\\_2](https://doi.org/10.1207/s15327957pspr1001_2)
- Barrett, L. F. (2013). Psychological construction: The darwinian approach to the science of emotion. *Emotion Review*, 5(4), 379–389. <https://doi.org/10.1177/1754073913489753>
- Barrett, L. F. (2017). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, 12(11), Article 1833. <https://doi.org/10.1093/scan/nsx060>
- Barrett, L. F. (2022). Context reconsidered: Complex signal ensembles, relational meaning, and population thinking in psychological science. *American Psychologist*, 77(8), 894–920. <https://doi.org/10.1037/amp0001054>
- Barrett, L. F., & Satpute, A. B. (2019). Historical pitfalls and new directions in the neuroscience of emotion. *Neuroscience Letters*, 693, 9–18. <https://doi.org/10.1016/j.neulet.2017.07.045>
- Barrett, L. F., & Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nature Reviews Neuroscience*, 16(7), 419–429. <https://doi.org/10.1038/nrn3950>
- Barsalou, L., & Wiemer-Hastings, K. (2005). Situating abstract concepts. In D. Pecher & R. A. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thought* (pp. 129–163). Cambridge University Press. <https://doi.org/10.1017/CBO9780511499968.007>
- Bartram, D. J. (1976). Levels of coding in picture-picture comparison tasks. *Memory & Cognition*, 4(5), 593–602. <https://doi.org/10.3758/BF03213223>
- Berthoz, S., Artiges, E., Van De Moortele, P.-F., Poline, J.-B., Rouquette, S., Consoli, S. M., & Martinot, J.-L. (2002). Effect of impaired recognition and expression of emotions on frontocingulate cortices: An fMRI study of men with alexithymia. *The American Journal of Psychiatry*, 159(6), 961–967. <https://doi.org/10.1176/appi.ajp.159.6.961>
- Bird, G., Silani, G., Brindley, R., White, S., Frith, U., & Singer, T. (2010). Empathic brain responses in insula are modulated by levels of alexithymia but not autism. *Brain: A Journal of Neurology*, 133(5), 1515–1525. <https://doi.org/10.1093/brain/awq060>
- Boden, M. T., & Thompson, R. J. (2017). Meta-analysis of the association between emotional clarity and attention to emotions. *Emotion Review*, 9(1), 79–85. <https://doi.org/10.1177/1754073915610640>
- Booth, R. D. L., & Happé, F. G. E. (2018). Evidence of reduced global processing in autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 48(4), 1397–1408. <https://doi.org/10.1007/s10803-016-2724-6>

- Borghi, A. M., Binkofski, F., Castelfranchi, C., Cimatti, F., Scorolli, C., & Tummolini, L. (2017). The challenge of abstract concepts. *Psychological Bulletin*, 143(3), 263–292. <https://doi.org/10.1037/bul0000089>
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181–204. <https://doi.org/10.1017/S0140525X12000477>
- Coan, J. A. (2010). Emergent ghosts of the emotion machine. *Emotion Review*, 2(3), 274–285. <https://doi.org/10.1177/1754073910361978>
- Cook, R., Brewer, R., Shah, P., & Bird, G. (2013). Alexithymia, not autism, predicts poor recognition of emotional facial expressions. *Psychological Science*, 24(5), 723–732. <https://doi.org/10.1177/0956797612463582>
- Curran, P. J., & Hussong, A. M. (2009). Integrative data analysis: The simultaneous analysis of multiple data sets. *Psychological Methods*, 14(2), 81–100. <https://doi.org/10.1037/a0015914>
- Dejonckheere, E., Mestdagh, M., Houben, M., Rutten, I., Sels, L., Kuppens, P., & Tuerlinckx, F. (2019). Complex affect dynamics add limited information to the prediction of psychological well-being. *Nature Human Behaviour*, 3(5), 478–491. <https://doi.org/10.1038/s41562-019-0555-0>
- Demenescu, L. R., Kortekaas, R., den Boer, J. A., & Aleman, A. (2010). Impaired attribution of emotion to facial expressions in anxiety and major depression. *PLOS ONE*, 5(12), Article e15058. <https://doi.org/10.1371/journal.pone.0015058>
- Díaz, R., & Prinz, J. (2023). The role of emotional awareness in evaluative judgment: Evidence from alexithymia. *Scientific Reports*, 13(1), Article 5183. <https://doi.org/10.1038/s41598-023-32242-y>
- Doyle, C. M., Lane, S. T., Brooks, J. A., Wilkins, R. W., Gates, K. M., & Lindquist, K. A. (2022). Unsupervised classification reveals consistency and degeneracy in neural network patterns of emotion. *Social Cognitive and Affective Neuroscience*, 17(11), 995–1006. <https://doi.org/10.1093/scan/nsac028>
- Dunsmoor, J. E., & Paz, R. (2015). Fear generalization and anxiety: Behavioral and neural mechanisms. *Biological Psychiatry*, 78(5), 336–343. <https://doi.org/10.1016/j.biopsych.2015.04.010>
- Eichmann, M., Kugel, H., & Suslow, T. (2008). Difficulty identifying feelings and automatic activation in the fusiform gyrus in response to facial emotion. *Perceptual and Motor Skills*, 107(3), 915–922. <https://doi.org/10.2466/pms.107.3.915-922>
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3–4), 169–200. <https://doi.org/10.1080/0269939208411068>
- Erbas, Y., Sels, L., Ceulemans, E., & Kuppens, P. (2016). Feeling me, feeling you: The relation between emotion differentiation and empathic accuracy. *Social Psychological and Personality Science*, 7(3), 240–247. <https://doi.org/10.1177/1948550616633504>
- Gendron, M. (2017). Revisiting diversity: Cultural variation reveals the constructed nature of emotion perception. *Current Opinion in Psychology*, 17, 145–150. <https://doi.org/10.1016/j.copsyc.2017.07.014>
- Grynbarg, D., Chang, B., Corneille, O., Maurage, P., Vermeulen, N., Berthoz, S., & Luminet, O. (2012). Alexithymia and the processing of emotional facial expressions (EEFs): Systematic review, unanswered questions and further perspectives. *PLOS ONE*, 7(8), Article e42429. <https://doi.org/10.1371/journal.pone.0042429>
- Habibi Asgarabad, M., Salehi Yegaei, P., Jafari, F., Azami-Aghdash, S., & Lumley, M. A. (2023). The relationship of alexithymia to pain and other symptoms in fibromyalgia: A systematic review and meta-analysis. *European Journal of Pain*, 27(3), 321–337. <https://doi.org/10.1002/ejp.2064>
- Hamaker, E. L., Ceulemans, E., Grasman, R. P. P. P., & Tuerlinckx, F. (2015). Modeling affect dynamics: State of the art and future challenges. *Emotion Review*, 7(4), 316–322. <https://doi.org/10.1177/1754073915590619>
- Hariri, A. R., Tessitore, A., Mattay, V. S., Fera, F., & Weinberger, D. R. (2002). The amygdala response to emotional stimuli: A comparison of faces and scenes. *NeuroImage*, 17(1), 317–323. <https://doi.org/10.1006/nimg.2002.1179>
- Hegefeld, H. M., Satpute, A. B., Ochsner, K. N., Davidow, J. Y., & Nook, E. C. (2023). Fluency generating emotion words correlates with verbal measures but not emotion regulation, alexithymia, or depressive symptoms. *Emotion*, 23(8), 2259–2269. <https://doi.org/10.1037/emo0001229>
- Hemming, L., Taylor, P., Haddock, G., Shaw, J., & Pratt, D. (2019). A systematic review and meta-analysis of the association between alexithymia and suicide ideation and behaviour. *Journal of Affective Disorders*, 254, 34–48. <https://doi.org/10.1016/j.jad.2019.05.013>
- Hobson, H., Brewer, R., Catmur, C., & Bird, G. (2019). The role of language in alexithymia: Moving towards a multiroute model of alexithymia. *Emotion Review*, 11(3), 247–261. <https://doi.org/10.1177/1754073919838528>
- Hoemann, K., Devlin, M., & Barrett, L. F. (2020). Comment: Emotions are abstract, conceptual categories that are learned by a predicting brain. *Emotion Review*, 12(4), 253–255. <https://doi.org/10.1177/1754073919897296>
- Hoemann, K., Nielson, C., Yuen, A., Gurera, J. W., Quigley, K. S., & Barrett, L. F. (2021). Expertise in emotion: A scoping review and unifying framework for individual differences in the mental representation of emotional experience. *Psychological Bulletin*, 147(11), 1159–1183. <https://doi.org/10.1037/bul0000327>
- Honkalahti, K., Hintikka, J., Antikainen, R., Lehtonen, J., & Viinamäki, H. (2001). Alexithymia in patients with major depressive disorder and comorbid cluster C personality disorders: A 6-month follow-up study. *Journal of Personality Disorders*, 15(3), 245–254. <https://doi.org/10.1521/pedi.15.3.245.19211>
- Honkalahti, K., Jokela, M., Lehto, S. M., Kivimäki, M., & Virtanen, M. (2022). Association between alexithymia and substance use: A systematic review and meta-analysis. *Scandinavian Journal of Psychology*, 63(5), 427–438. <https://doi.org/10.1111/sjop.12821>
- Honkalahti, K., Saarinen, P., Hintikka, J., Virtanen, V., & Viinamäki, H. (1999). Factors associated with alexithymia in patients suffering from depression. *Psychotherapy and Psychosomatics*, 68(5), 270–275. <https://doi.org/10.1159/000012343>
- Houben, M., Van Den Noortgate, W., & Kuppens, P. (2015). The relation between short-term emotion dynamics and psychological well-being: A meta-analysis. *Psychological Bulletin*, 141(4), 901–930. <https://doi.org/10.1037/a0038822>
- Huff, C., & Tingley, D. (2015). “Who are these people?” Evaluating the demographic characteristics and political preferences of MTurk survey respondents. *Research & Politics*, 2(3), 1–12. <https://doi.org/10.1177/2053168015604648>
- Hutchinson, J. B., & Barrett, L. F. (2019). The power of predictions: An emerging paradigm for psychological research. *Current Directions in Psychological Science*, 28(3), 280–291. <https://doi.org/10.1177/0963721419831992>
- Iigaya, K., Yi, S., Wahle, I. A., Tanwisuth, K., & O’Doherty, J. P. (2021). Aesthetic preference for art can be predicted from a mixture of low- and high-level visual features. *Nature Human Behaviour*, 5(6), 743–755. <https://doi.org/10.1038/s41562-021-01124-6>
- Israelashvili, J., Oosterwijk, S., Sauter, D., & Fischer, A. (2019). Knowing me, knowing you: Emotion differentiation in oneself is associated with recognition of others’ emotions. *Cognition and Emotion*, 33(7), 1461–1471. <https://doi.org/10.1080/02699312019.1577221>
- Itkes, O., Kimchi, R., Haj-Ali, H., Shapiro, A., & Kron, A. (2017). Dissociating affective and semantic valence. *Journal of Experimental Psychology: General*, 146(7), 924–942. <https://doi.org/10.1037/xge0000291>
- Itkes, O., & Kron, A. (2019). Affective and semantic representations of valence: A conceptual framework. *Emotion Review*, 11(4), 283–293. <https://doi.org/10.1177/1754073919868759>
- Kano, M., Fukudo, S., Gyoba, J., Kamachi, M., Tagawa, M., Mochizuki, H., Itoh, M., Hongo, M., & Yanai, K. (2003). Specific brain processing of facial expressions in people with alexithymia: An H2 15O-PET study.

- Brain: A Journal of Neurology*, 126(6), 1474–1484. <https://doi.org/10.1093/brain/awg131>
- Kano, M., Hamaguchi, T., Itoh, M., Yanai, K., & Fukudo, S. (2007). Correlation between alexithymia and hypersensitivity to visceral stimulation in human. *Pain*, 132(3), 252–263. <https://doi.org/10.1016/j.pain.2007.01.032>
- Keltner, D., Sauter, D., Tracy, J., & Cowen, A. (2019). Emotional expression: Advances in basic emotion theory. *Journal of Nonverbal Behavior*, 43(2), 133–160. <https://doi.org/10.1007/s10919-019-00293-3>
- Kim, J. H., Lee, S. J., Rim, H. D., Kim, H. W., Bae, G. Y., & Chang, S. M. (2008). The relationship between alexithymia and general symptoms of patients with depressive disorders. *Psychiatry Investigation*, 5(3), 179–185. <https://doi.org/10.4306/pi.2008.5.3.179>
- Kinnaird, E., Stewart, C., & Tchanturia, K. (2019). Investigating alexithymia in autism: A systematic review and meta-analysis. *European Psychiatry*, 55, 80–89. <https://doi.org/10.1016/j.eurpsy.2018.09.004>
- Kugel, H., Eichmann, M., Dannlowski, U., Ohrmann, P., Bauer, J., Arrolt, V., Heindel, W., & Suslow, T. (2008). Alexithymic features and automatic amygdala reactivity to facial emotion. *Neuroscience Letters*, 435(1), 40–44. <https://doi.org/10.1016/j.neulet.2008.02.005>
- Kuppens, P., & Verduyn, P. (2017). Emotion dynamics. *Current Opinion in Psychology*, 17, 22–26. <https://doi.org/10.1016/j.copsyc.2017.06.004>
- Lane, R. D., Quinlan, D. M., Schwartz, G. E., Walker, P. A., & Zeitlin, S. B. (1990). The Levels of Emotional Awareness Scale: A cognitive-developmental measure of emotion. *Journal of Personality Assessment*, 55(1–2), 124–134. <https://doi.org/10.1080/00223891.1990.9674052>
- Lane, R. D., Sechrest, L., Riedel, R., Shapiro, D. E., & Kaszniak, A. W. (2000). Pervasive emotion recognition deficit common to alexithymia and the repressive coping style. *Psychosomatic Medicine*, 62(4), 492–501. <https://doi.org/10.1097/000006842-200007000-00007>
- Lane, R. D., Weihs, K. L., Herring, A., Hishaw, A., & Smith, R. (2015). Affective agnosia: Expansion of the alexithymia construct and a new opportunity to integrate and extend Freud's legacy. *Neuroscience and Biobehavioral Reviews*, 55, 594–611. <https://doi.org/10.1016/j.neubiorev.2015.06.007>
- Larsen, J. T., Berntson, G. G., Poehlmann, K. M., Ito, T. A., & Cacioppo, J. T. (2008). The psychophysiology of emotion. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 180–195). Guilford Press.
- Lee, B.-T., Lee, H.-Y., Park, S.-A., Lim, J.-Y., Tae, W. S., Lee, M.-S., Joe, S.-H., Jung, I.-K., & Ham, B.-J. (2011). Neural substrates of affective face recognition in alexithymia: A functional magnetic resonance imaging study. *Neuropsychobiology*, 63(2), 119–124. <https://doi.org/10.1159/000318086>
- Lee, K. M., Ferreira-Santos, F., & Satpute, A. B. (2021). Predictive processing models and affective neuroscience. *Neuroscience and Biobehavioral Reviews*, 131, 211–228. <https://doi.org/10.1016/j.neubiorev.2021.09.009>
- Lee, K. M., Lindquist, K. A., Arbuckle, N. L., Mowrer, S. M., & Payne, B. K. (2020). An indirect measure of discrete emotions. *Emotion*, 20(4), 659–676. <https://doi.org/10.1037/emo0000577>
- Leising, D., Grande, T., & Faber, R. (2009). The Toronto Alexithymia Scale (TAS-20): A measure of general psychological distress. *Journal of Research in Personality*, 43(4), 707–710. <https://doi.org/10.1016/j.jrp.2009.03.009>
- Levenson, R. W. (2003). Autonomic specificity and emotion. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 212–224). Oxford University Press.
- Leweke, F., Leichsenring, F., Kruse, J., & Hermes, S. (2011). Is alexithymia associated with specific mental disorders? *Psychopathology*, 45(1), 22–28. <https://doi.org/10.1159/000325170>
- Lieberman, M. D., Eisenberger, N. I., Crockett, M. J., Tom, S. M., Pfeifer, J. H., & Way, B. M. (2007). Putting feelings into words. *Psychological Science*, 18(5), 421–428. <https://doi.org/10.1111/j.1467-9280.2007.01916.x>
- Lindquist, K. A., & Barrett, L. F. (2008). Emotional complexity. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 513–530). Guilford Press.
- Lindquist, K. A., Jackson, J. C., Leshin, J., Satpute, A. B., & Gendron, M. (2022). The cultural evolution of emotion. *Nature Reviews Psychology*, 1(11), 669–681. <https://doi.org/10.1038/s44159-022-00105-4>
- Lindquist, K. A., MacCormack, J. K., & Shablack, H. (2015). The role of language in emotion: Predictions from psychological constructionism. *Frontiers in Psychology*, 6, Article 444. <https://doi.org/10.3389/fpsyg.2015.00444>
- Lindquist, K. A., Satpute, A. B., & Gendron, M. (2015). Does language do more than communicate emotion? *Current Directions in Psychological Science*, 24(2), 99–108. <https://doi.org/10.1177/0963721414553440>
- Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences*, 35(3), 121–143. <https://doi.org/10.1017/S0140525X11000446>
- Mantani, T., Okamoto, Y., Shirao, N., Okada, G., & Yamawaki, S. (2005). Reduced activation of posterior cingulate cortex during imagery in subjects with high degrees of alexithymia: A functional magnetic resonance imaging study. *Biological Psychiatry*, 57(9), 982–990. <https://doi.org/10.1016/j.biopsych.2005.01.047>
- Marchesi, C., Ossola, P., Tonna, M., & De Panfilis, C. (2014). The TAS-20 more likely measures negative affects rather than alexithymia itself in patients with major depression, panic disorder, eating disorders and substance use disorders. *Comprehensive Psychiatry*, 55(4), 972–978. <https://doi.org/10.1016/j.comppsych.2013.12.008>
- Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and Emotion*, 23(2), 209–237. <https://doi.org/10.1080/02699930802204677>
- Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT V2.0. *Emotion*, 3(1), 97–105. <https://doi.org/10.1037/1528-3542.3.1.97>
- McVeigh, K., Kleckner, I. R., Quigley, K. S., & Satpute, A. B. (2024). Fear-related psychophysiological patterns are situation and individual dependent: A Bayesian model comparison approach. *Emotion*, 24(2), 506–521. <https://doi.org/10.1037/emo0001265>
- Mériau, K., Wartenburger, I., Kazzer, P., Prehn, K., Lammers, C.-H., van der Meer, E., Villringer, A., & Heekeren, H. R. (2006). A neural network reflecting individual differences in cognitive processing of emotions during perceptual decision making. *NeuroImage*, 33(3), 1016–1027. <https://doi.org/10.1016/j.neuroimage.2006.07.031>
- Mikels, J. A., & Reuter-Lorenz, P. A. (2019). Affective working memory: An integrative psychological construct. *Perspectives on Psychological Science*, 14(4), 543–559. <https://doi.org/10.1177/1745691619837597>
- Morgan, J. K., Izard, C. E., & King, K. A. (2010). Construct validity of the emotion matching task: Preliminary evidence for convergent and criterion validity of a new emotion knowledge measure for young children. *Social Development*, 19(1), 52–70. <https://doi.org/10.1111/j.1467-9507.2008.00529.x>
- Moriguchi, Y., & Komaki, G. (2013). Neuroimaging studies of alexithymia: Physical, affective, and social perspectives. *BioPsychoSocial Medicine*, 7(1), Article 8. <https://doi.org/10.1186/1751-0759-7-8>
- Murphy, G. L. (2002). *The big book of concepts*. The MIT Press. <https://doi.org/10.7551/mitpress/1602.001.0001>
- Nook, E. C., Lindquist, K. A., & Zaki, J. (2015). A new look at emotion perception: Concepts speed and shape facial emotion recognition. *Emotion*, 15(5), 569–578. <https://doi.org/10.1037/a0039166>
- Nook, E. C., Sasse, S. F., Lambert, H. K., McLaughlin, K. A., & Somerville, L. H. (2017). Increasing verbal knowledge mediates development of multidimensional emotion representations. *Nature Human Behaviour*, 1(12), 881–889. <https://doi.org/10.1038/s41562-017-0238-7>
- Nook, E. C., Satpute, A. B., & Ochsner, K. N. (2021). Emotion naming impedes both cognitive reappraisal and mindful acceptance strategies of

- emotion regulation. *Affective Science*, 2(2), 187–198. <https://doi.org/10.1007/s42761-021-00036-y>
- Nook, E. C., Stavish, C. M., Sasse, S. F., Lambert, H. K., Mair, P., McLaughlin, K. A., & Somerville, L. H. (2020). Charting the development of emotion comprehension and abstraction from childhood to adulthood using observer-rated and linguistic measures. *Emotion*, 20(5), 773–792. <https://doi.org/10.1037/emo0000609>
- Ogren, M., & Johnson, S. P. (2020). Intermodal emotion matching at 15 months, but not 9 or 21 months, predicts early childhood emotion understanding: A longitudinal investigation. *Cognition and Emotion*, 34(7), 1343–1356. <https://doi.org/10.1080/02699931.2020.1743236>
- Panksepp, J. (2011). The basic emotional circuits of mammalian brains: Do animals have affective lives? *Neuroscience and Biobehavioral Reviews*, 35(9), 1791–1804. <https://doi.org/10.1016/j.neubiorev.2011.08.003>
- Parker, P. D., Prkachin, K. M., & Prkachin, G. C. (2005). Processing of facial expressions of negative emotion in alexithymia: The influence of temporal constraint. *Journal of Personality*, 73(4), 1087–1107. <https://doi.org/10.1111/j.1467-6494.2005.00339.x>
- Poquérusse, J., Pastore, L., Dellantonio, S., & Esposito, G. (2018). Alexithymia and autism spectrum disorder: A complex relationship. *Frontiers in Psychology*, 9, Article 1196. <https://doi.org/10.3389/fpsyg.2018.01196>
- Preece, D. A., Becerra, R., Allan, A., Robinson, K., & Dandy, J. (2017). Establishing the theoretical components of alexithymia via factor analysis: Introduction and validation of the attention-appraisal model of alexithymia. *Personality and Individual Differences*, 119, 341–352. <https://doi.org/10.1016/j.paid.2017.08.003>
- Preece, D. A., Becerra, R., Boyes, M. E., Northcott, C., McGillivray, L., & Hasking, P. A. (2020). Do self-report measures of alexithymia measure alexithymia or general psychological distress? A factor analytic examination across five samples. *Personality and Individual Differences*, 155, Article 109721. <https://doi.org/10.1016/j.paid.2019.109721>
- Preece, D. A., Becerra, R., Robinson, K., Dandy, J., & Allan, A. (2018). The psychometric assessment of alexithymia: Development and validation of the Perth Alexithymia Questionnaire. *Personality and Individual Differences*, 132, 32–44. <https://doi.org/10.1016/j.paid.2018.05.011>
- Preece, D. A., & Gross, J. J. (2023). Conceptualizing alexithymia. *Personality and Individual Differences*, 215, Article 112375. <https://doi.org/10.1016/j.paid.2023.112375>
- Prkachin, G. C., Casey, C., & Prkachin, K. M. (2009). Alexithymia and perception of facial expressions of emotion. *Personality and Individual Differences*, 46(4), 412–417. <https://doi.org/10.1016/j.paid.2008.11.010>
- Quoidbach, J., Gruber, J., Mikolajczak, M., Kogan, A., Kotsou, I., & Norton, M. I. (2014). Emodiversity and the emotional ecosystem. *Journal of Experimental Psychology: General*, 143(6), 2057–2066. <https://doi.org/10.1037/a0038025>
- Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1(3), 385–401. <https://doi.org/10.1177/014662167700100306>
- Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, 114(3), 510–532. <https://doi.org/10.1037/0033-295X.114.3.510>
- Ruba, A. L., Meltzoff, A. N., & Repacholi, B. M. (2021). Linguistic and developmental influences on superordinate facial configuration categorization in infancy. *Infancy*, 26(6), 857–876. <https://doi.org/10.1111/infra.12430>
- Saarimäki, H., Gotsopoulos, A., Jääskeläinen, I. P., Lampinen, J., Vuilleumier, P., Hari, R., Sams, M., & Nummenmaa, L. (2016). Discrete neural signatures of basic emotions. *Cerebral Cortex*, 26(6), 2563–2573. <https://doi.org/10.1093/cercor/bhw086>
- Satpute, A. B., Badre, D., & Ochsner, K. N. (2014). Distinct regions of prefrontal cortex are associated with the controlled retrieval and selection of social information. *Cerebral Cortex*, 24(5), 1269–1277. <https://doi.org/10.1093/cercor/bhs408>
- Satpute, A. B., & Lindquist, K. A. (2019). The default mode network's role in discrete emotion. *Trends in Cognitive Sciences*, 23(10), 851–864. <https://doi.org/10.1016/j.tics.2019.07.003>
- Satpute, A. B., & Lindquist, K. A. (2021). At the neural intersection between language and emotion. *Affective Science*, 2(2), 207–220. <https://doi.org/10.1007/s42761-021-00032-2>
- Satpute, A. B., Shu, J., Weber, J., Roy, M., & Ochsner, K. N. (2013). The functional neural architecture of self-reports of affective experience. *Biological Psychiatry*, 73(7), 631–638. <https://doi.org/10.1016/j.biopsych.2012.10.001>
- Schroeders, U., Kubera, F., & Gnambs, T. (2022). The structure of the Toronto Alexithymia Scale (TAS-20): A meta-analytic confirmatory factor analysis. *Assessment*, 29(8), 1806–1823. <https://doi.org/10.1177/10731911211033894>
- Schweizer, S., Grahn, J., Hampshire, A., Mobbs, D., & Dalgleish, T. (2013). Training the emotional brain: Improving affective control through emotional working memory training. *The Journal of Neuroscience*, 33(12), 5301–5311. <https://doi.org/10.1523/JNEUROSCI.2593-12.2013>
- Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends in Cognitive Sciences*, 17(11), 565–573. <https://doi.org/10.1016/j.tics.2013.09.007>
- Siegel, E. H., Sands, M. K., Van den Noortgate, W., Condon, P., Chang, Y., Dy, J., Quigley, K. S., & Barrett, L. F. (2018). Emotion fingerprints or emotion populations? A meta-analytic investigation of autonomic features of emotion categories. *Psychological Bulletin*, 144(4), 343–393. <https://doi.org/10.1037/bul0000128>
- Sifneos, P. E. (1973). The prevalence of “alexithymic” characteristics in psychosomatic patients. *Psychotherapy and Psychosomatics*, 22(2–6), 255–262. <https://doi.org/10.1159/000286529>
- Sifneos, P. E. (1996). Alexithymia: Past and present. *The American Journal of Psychiatry*, 153(7), 137–142. <https://doi.org/10.1176/appi.153.7.137>
- Skerry, A. E., & Saxe, R. (2014). A common neural code for perceived and inferred emotion. *The Journal of Neuroscience*, 34(48), 15997–16008. <https://doi.org/10.1523/JNEUROSCI.1676-14.2014>
- Smidt, K. E., & Suvak, M. K. (2015). A brief, but nuanced, review of emotional granularity and emotion differentiation research. *Current Opinion in Psychology*, 3, 48–51. <https://doi.org/10.1016/j.copsyc.2015.02.007>
- Smith, R., Killgore, W. D. S., & Lane, R. D. (2018). The structure of emotional experience and its relation to trait emotional awareness: A theoretical review. *Emotion*, 18(5), 670–692. <https://doi.org/10.1037/emo0000376>
- Smith, R., Lane, R. D., Parr, T., & Friston, K. J. (2019). Neurocomputational mechanisms underlying emotional awareness: Insights afforded by deep active inference and their potential clinical relevance. *Neuroscience and Biobehavioral Reviews*, 107, 473–491. <https://doi.org/10.1016/j.neubiorev.2019.09.002>
- Smith, R., Parr, T., & Friston, K. J. (2019). Simulating emotions: An active inference model of emotional state inference and emotion concept learning. *Frontiers in Psychology*, 10, Article 2844. <https://doi.org/10.3389/fpsyg.2019.02844>
- Spunt, R. P., & Adolphs, R. (2017). A new look at domain specificity: Insights from social neuroscience. *Nature Reviews Neuroscience*, 18(9), 559–567. <https://doi.org/10.1038/nrn.2017.76>
- Spunt, R. P., & Adolphs, R. (2019). The neuroscience of understanding the emotions of others. *Neuroscience Letters*, 693, 44–48. <https://doi.org/10.1016/j.neulet.2017.06.018>
- Suvak, M. K., Litz, B. T., Sloan, D. M., Zanarini, M. C., Barrett, L. F., & Hofmann, S. G. (2011). Emotional granularity and borderline personality disorder. *Journal of Abnormal Psychology*, 120(2), 414–426. <https://doi.org/10.1037/a0021808>
- Taylor, G. J., & Bagby, R. M. (2004). New trends in alexithymia research. *Psychotherapy and Psychosomatics*, 73(2), 68–77. <https://doi.org/10.1159/000075537>

- Taylor, G. J., Bagby, R. M., & Parker, J. D. (1999). *Disorders of affect regulation: Alexithymia in medical and psychiatric illness*. Cambridge University Press.
- Taylor, G. J., Bagby, R. M., & Parker, J. D. A. (2016). What's in the name "alexithymia"? A commentary on "affective agnosia: Expansion of the alexithymia construct and a new opportunity to integrate and extend Freud's legacy". *Neuroscience and Biobehavioral Reviews*, 68, 1006–1020. <https://doi.org/10.1016/j.neubiorev.2016.05.025>
- Thompson, R. J., Dizén, M., & Berenbaum, H. (2009). The unique relations between emotional awareness and facets of affective instability. *Journal of Research in Personality*, 43(5), 875–879. <https://doi.org/10.1016/j.jrp.2009.07.006>
- Tugade, M. M., Fredrickson, B. L., & Feldman Barrett, L. (2004). Psychological resilience and positive emotional granularity: Examining the benefits of positive emotions on coping and health. *Journal of Personality*, 72(6), 1161–1190. <https://doi.org/10.1111/j.1467-6494.2004.00294.x>
- Uljarevic, M., & Hamilton, A. (2013). Recognition of emotions in autism: A formal meta-analysis. *Journal of Autism and Developmental Disorders*, 43(7), 1517–1526. <https://doi.org/10.1007/s10803-012-1695-5>
- van der Velde, J., Servaas, M. N., Goerlich, K. S., Bruggeman, R., Horton, P., Costafreda, S. G., & Aleman, A. (2013). Neural correlates of alexithymia: A meta-analysis of emotion processing studies. *Neuroscience and Biobehavioral Reviews*, 37(8), 1774–1785. <https://doi.org/10.1016/j.neubiorev.2013.07.008>
- Wang, Y., Kragel, P. A., & Satpute, A. B. (2024). Neural predictors of fear depend on the situation. *The Journal of Neuroscience*, 44(46), Article e0142232024. <https://doi.org/10.1523/JNEUROSCI.0142-23.2024>
- Waugh, C. E., Lemus, M. G., & Gotlib, I. H. (2014). The role of the medial frontal cortex in the maintenance of emotional states. *Social Cognitive and Affective Neuroscience*, 9(12), 2001–2009. <https://doi.org/10.1093/scan/nsu011>
- Weissman, D. G., Nook, E. C., Dews, A. A., Miller, A. B., Lambert, H. K., Sasse, S. F., Somerville, L. H., & McLaughlin, K. A. (2020). Low emotional awareness as a transdiagnostic mechanism underlying psychopathology in adolescence. *Clinical Psychological Science*, 8(6), 971–988. <https://doi.org/10.1177/2167702620923649>
- Widen, S. C., Pochedly, J. T., & Russell, J. A. (2015). The development of emotion concepts: A story superiority effect in older children and adolescents. *Journal of Experimental Child Psychology*, 131, 186–192. <https://doi.org/10.1016/j.jecp.2014.10.009>
- Widen, S. C., & Russell, J. A. (2008). Young children's understanding of other's emotions. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., Vol. 3, pp. 348–363). Guilford Press.
- Wilson-Mendenhall, C. D., Barrett, L. F., Simmons, W. K., & Barsalou, L. W. (2011). Grounding emotion in situated conceptualization. *Neuropsychologia*, 49(5), 1105–1127. <https://doi.org/10.1016/j.neuropsychologia.2010.12.032>
- Zahid, A., Taylor, G. J., Lau, S. C. L., Stone, S., & Bagby, R. M. (2024). Examining the incremental validity of the Perth Alexithymia Questionnaire (PAQ) relative to the 20-item Toronto Alexithymia Scale (TAS-20). *Journal of Personality Assessment*, 106(2), 242–253. <https://doi.org/10.1080/0023891.2023.2201831>

Received May 25, 2023

Revision received November 26, 2024

Accepted December 7, 2024 ■