

RESEARCH ARTICLE

Basic emotion recognition of children on the autism spectrum is enhanced in music and typical for faces and voices

Shalini Sivathasan^{1,2,3}, Hadas Dahary^{1,2,3}, Jacob A. Burack^{1,2}, Eve-Marie Quintin^{1,2,3*}

1 Department of Educational and Counselling Psychology, McGill University, Montreal, Quebec, Canada, **2** Azrieli Centre for Autism Research, Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada, **3** Center for Research on Music, Brain, and Language, McGill University, Montreal, Quebec, Canada

* eve-marie.quintin@mcgill.ca

OPEN ACCESS

Citation: Sivathasan S, Dahary H, Burack JA, Quintin E-M (2023) Basic emotion recognition of children on the autism spectrum is enhanced in music and typical for faces and voices. PLoS ONE 18(1): e0279002. <https://doi.org/10.1371/journal.pone.0279002>

Editor: Lucia Billeci, Istituto di Fisiologia Clinica Consiglio Nazionale delle Ricerche, ITALY

Received: June 22, 2022

Accepted: November 28, 2022

Published: January 11, 2023

Copyright: © 2023 Sivathasan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Due to the sensitive and identifying nature of these locally collected human research participant data, these data are not publicly available. All relevant data are within the manuscript and its [Supporting information files](#). Questions about these data can be directed to the Associate Director of Research Ethics, Office of Research Ethics and Compliance, McGill University, lynda.mcneil@mcgill.ca.

Funding: Support for this research was provided by the Social Sciences and Humanities Research

Abstract

In contrast with findings of reduced facial and vocal emotional recognition (ER) accuracy, children on the autism spectrum (AS) demonstrate comparable ER skills to those of typically-developing (TD) children using music. To understand the specificity of purported ER differences, the goal of this study was to examine ER from music compared with faces and voices among children on the AS and TD children. Twenty-five children on the AS and 23 TD children (6–13 years) completed an ER task, using categorical (happy, sad, fear) and dimensional (valence, arousal) ratings, of emotions presented via music, faces, or voices. Compared to the TD group, the AS group showed a relative ER strength from music, and comparable performance from faces and voices. Although both groups demonstrated greater vocal ER accuracy, the children on the AS performed equally well with music and faces, whereas the TD children performed better with faces than with music. Both groups performed comparably with dimensional ratings, except for greater variability by the children on the AS in valence ratings for happy emotions. These findings highlight a need to re-examine ER of children on the AS, and to consider how facilitating strengths-based approaches can re-shape our thinking about and support for persons on the AS.

Introduction

Differing patterns of emotion recognition (ER) have long been cited as significant components of the socialization profile that is characteristic of persons on the autism spectrum (AS) as compared “typically-developing” (TD), persons [1–7]. However, virtually all of the evidence of relative ER challenges or “deficits” has been based on findings with paradigms in which the presentation of emotions occurs in the context of socially explicit facial and vocal stimuli. This evidence of reduced ER accuracy diverges from that of comparable abilities between persons on the AS and TD persons to recognize emotions when the tasks are presented in the context of music [8–10], a domain in which persons on the AS have been found to display a variety of unique abilities [e.g., 11–13]. To better understand the nature of these disparate findings, we

Council of Canada (<https://www.sshrc-crsh.gc.ca/>) to SS and EMQ, Fonds de Recherche du Québec – Santé (<https://frq.gouv.qc.ca/en/health/>) to EMQ, and the Organization for Autism Research (<https://researchautism.org/>) to SS. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

investigated ER skills with musical, as compared to facial and vocal, stimuli using both categorical and dimensional response options in order both to provide a nuanced understanding of ER processing across various types of stimuli and to consider the ways that strength-based approaches can re-shape our thinking about persons on the AS.

Music is a stimulus with which persons on the AS regularly and readily engage in their day-to-day lives, and with which they have demonstrated a variety of skills and strengths [13]. Persons on the AS have been found to show enhanced pitch discrimination of musical tones and melodies [14–18] and musical memory [19–21] as compared to mental age (MA)-matched TD persons. Among children on the AS, verbal production may be enhanced through the incorporation of music into speech- and language-focused interventions, particularly for children who have greater language and learning needs [e.g., 22–25]. Further, music therapy has been shown to improve social interaction and communication skills, such as joint attention [26, 27], turn taking [26], and social communication [28] when the development of these skills is the target of the intervention.

The study of ER skills of persons on the AS involves a similar strengths-based approach via the inclusion of musical stimuli. For example, Heaton et al. [8] found that children on the AS show comparable abilities to MA-matched TD children in distinguishing music conveying happy (typically in a major musical scale) and sad (in a minor scale) emotions. Concordantly, comparable ER ratings of happy, sad, and fearful musical excerpts between children and adolescents on the AS and their TD peers have been found when differences in verbal IQ were controlled [9, 10]. Similar performance on behavioral ER tasks as well as on those involving the activation of emotion processing and reward neurocircuitry when listening to happy and sad music have been shown between persons on the AS and TD persons [29, 30]. The evidence from these musical ER studies suggests that persons on the AS appear to be able to recognize basic emotions conveyed by music comparably to their TD peers of a similar developmental level. In contrast, the findings from meta-analyses of behavioral ER studies indicate medium to large overall effect sizes across studies indicating lower accuracy levels on recognition tasks of basic emotions among adults and children on the AS as compared to TD participants when socially explicit stimuli (e.g., faces, voices) were used [6, 7, 31]. This suggests a discrepancy between findings based on socially explicit stimuli versus less socially explicit stimuli that can also convey emotions (i.e., music).

Attempts to reconcile such discrepancies need to address how music differs in its conveyance of emotion as compared to explicitly social facial and vocal stimuli. Emotions are most commonly presented in an interpersonal or self-reflective context, such as through social observation and interaction or through an intrapersonal emotional experience. This use of socially explicit facial and vocal stimuli in ER studies with persons on the AS may contribute to the finding of reduced ER accuracy on behavioral tasks. For example, differences in facial ER accuracy among persons on the AS relative to TD persons could be related to early developmental differences in reduced attention to eyes and faces [32] or to the reduced value and salience of social versus non-social stimuli that begins in childhood [33, 34]. In contrast, music can elicit an affective reaction in the listener without requiring a need to understand or empathize with the composer's mental states or emotional intentions [8, 35]. Emotions are conveyed in music through variations in structural auditory aspects such as tempo, timbre, pitch, and mode, and thereby afford a unique yet common and familiar vehicle for emotional experience [36]. Accordingly, the primary goal of the current study is to reconcile reported differences in ER among children on the AS, and specifically to determine whether music affords a relative ER strength in comparison with explicitly social (i.e., facial and vocal) stimuli, and whether such a pattern of strengths and challenges is observed among their TD peers. If children on the AS demonstrate relative ER strengths using less socially explicit stimuli and TD children

demonstrate relative ER strengths using more socially explicit stimuli, we could speculate that differing routes for understanding and teaching ER skills are available to children with diverse strengths and needs.

The disparities in findings across studies may also be due to the differences in the measurement techniques that are used in experimental designs and methodologies in ER research [e.g., 7, 37–40]. Based on the theory that basic emotions are universally expressed and recognized [41, 42], ER abilities are frequently measured with behavioral paradigms using these discrete categorical representations (e.g., happy, sad, fear) through a forced-choice response format with verbal labels. However, emotional expressions can also be considered a product of dimensional attributes, particularly valence (positive to negative) and arousal (high to low energy) [43], and therefore can also be measured using continuous ratings [43, 44; 36 for a review]. While some of the earliest experimental studies of ER among persons on the AS have involved a categorical approach [e.g., 4, 5], the use of dimensional methods is becoming increasingly more common [e.g., 29, 45]. Accordingly, in order to capture dynamic aspects of emotions across each domain along a continuum of valence and arousal ratings, our second goal was to provide a direct comparison of musical, facial, and vocal ER among children on the AS and TD children with dimensional ratings versus traditional categorical response options.

Objectives and hypotheses

The main objective of this study was to compare the processing of basic emotions using musical, facial, and vocal stimuli among children on the AS and TD children. In comparing the groups, we hypothesized that the AS group would demonstrate a relative strength in musical ER compared to the TD group, whereas the TD group might display a strength in ER with socially explicit stimuli (faces, voices) relative to the AS group. We also hypothesized that there would be discrepancies in ER with a relative weakness in facial and vocal ER as compared to musical ER within the AS group, whereas we did not expect any significant difference in performance among musical, facial, and vocal ER within the TD group.

We also explored the use of dimensional ratings of valence and arousal across the musical, facial, and vocal stimulus types. As this aim was exploratory, we did not venture specific hypotheses regarding between or within group differences. To allow for a direct comparison of ER across the three domains, we developed an integrated experimental task displaying the musical, facial, and vocal stimuli, with categorical and dimensional response options presented for each item presented. We focused on comparisons from the three basic emotions of happy, sad, and fear that can be reliably conveyed by music [8–10].

Method

Participants

The initial group of participants included 29 children on the AS recruited from special education schools or classrooms and 24 TD children recruited from schools or the general community in Quebec, Canada. The data from five participants (4 AS and 1 TD) were excluded from the analyses because one participant on the AS did not complete the verbal comprehension portion of the cognitive test and the scores on the experimental task of the other four participants were multivariate outliers. Accordingly, the final group of participants ($N = 48$) included 25 children on the AS (19 boys) aged 9–13 years ($M = 11$ years) and 23 TD children (11 boys) aged 6–12 years ($M = 9.7$ years).

All of the participants on the AS had an educational code of autism spectrum disorder derived from expert diagnoses from pediatricians, child psychiatrists, or psychologists (for a discussion of the educational code system in Quebec, see [46, 47]). The Parent and Teacher

Table 1. Participant characteristics of the AS and TD groups.

	AS Group (<i>n</i> = 25)			TD Group (<i>n</i> = 23)			<i>p</i>
	M	SD	Range	M	SD	Range	
Chronological Age	11.02	1.12	9–13	9.70	1.63	6–12	.003
Verbal Mental Age	8.48	1.62	5–11	10.03	2.56	5–15	.018
VCI	77	14	51–104	102	14	73–128	< .0001
FSIQ ^a	79	14	54–109	104	13	78–124	< .0001
SRS-2	71	9	54–88	49	8	40–62	< .0001

Note. Mean (M), standard deviation (SD), ranges, and *p* values of an independent sample *t*-test between the AS group and the TD group. Chronological age (CA), and Verbal Mental Age (VMA) are displayed in years. Verbal Comprehension Index (VCI) and Full-Scale IQ (FSIQ) = standard score. Social Responsiveness Scale-2 (SRS-2) Total Score = T-Score.

^aAS group *n* = 24, TD group *n* = 22.

<https://doi.org/10.1371/journal.pone.0279002.t001>

versions of the Social Responsiveness Scale, 2nd edition (SRS-2 [48]), were also used to ascertain the presence or absence of autism traits among the AS and TD groups, respectively. The average SRS Total T-Score was greater than the clinical cut-off of 60 for the AS group and below 60 for the TD group. Four participants on the AS had T-Scores slightly below (between 54–59) the cutoff and three TD participants had scores at or above (60–62). As a similar pattern of results were found when the analyses were run with and without these seven participants with the categorical response option, the full dataset was used in the reported analyses.

The verbal scales of the Wechsler Intelligence Scale for Children, Fifth Edition (WISC-V [49]) in English or French, or the Wechsler Abbreviated Intelligence Scale, Second Edition (WASI-II [50]) in English (no French version available) were used to estimate cognitive skills, and scores obtained were derived from the Verbal Comprehension Index (VCI) and Full-Scale IQ. Consistent with recommendations to consider the effects of verbal cognitive ability on ER task performance, and to account for variability in VCI and age ranges between groups [51–53], verbal mental age (VMA) was calculated using the participants' VCI and chronological age (AS group VMA: *M* = 8.48 [5–11 years]; TD group VMA: *M* = 10.03 [5–15 years]).

The AS and TD groups differed significantly in terms of SRS-2 Total Scores, chronological age, IQ, and VMA (Table 1). Thus, VMA was considered as a covariate in our analyses of the group comparisons. Maternal educational attainment levels and annual income reported by 75% of participants' caregivers did not differ significantly between groups (see S1 Table). Specific demographic data on race & ethnicity were not systematically collected during this study and were, therefore, not available for the analyses.

Experimental task

Musical, facial, and vocal stimuli. The ER task included 72 different musical, facial and vocal stimuli depicting emotions of happy, sad, or fear. The task was divided into 3 blocks (one per stimulus type: music, faces, voices), with each block containing 24 items (8 happy, 8 sad, and 8 fear). Each item within a block was presented briefly on the computer screen for 1.5–2 seconds. The vocal and musical stimuli selected from the validated sets described below were purposefully designed to be short emotional “bursts” representing the minimum basic emotional information required to appraise the emotion without providing extra semantic (or any linguistic) content [54, 55]. Selected facial stimuli were thus also presented at a short duration to match that of the auditory stimuli, as well as to mitigate against potentially making the task too easy by providing longer stimulus presentations [37]. The order of the three stimulus types

(music, faces, voices) was counterbalanced across the participants and the order of the items within each stimulus block was randomized.

Faces. The participants viewed 24 color photos of closed-mouth faces (happy, sad, or fear) selected from the NimStim Set of Facial Expressions [56], a well-validated set of photos of facial expressions conveying basic emotions. The photos were of Asian, Black, Latin-American, and White adults, providing a racially and ethnically diverse sample. In total, 8 unique adult actors ($n = 4$ women) represented 3 emotions (3 emotions \times 2 sexes \times 4 ethnicities = 24 faces). Each selected facial expression had a reliability coefficient of .78 or greater, based on validation with TD adults [54].

Voices. The participants listened to 24 selected vocal sounds from the Montreal Affective Voices battery (MAV [54]), a validated set of nonverbal emotional vocalizations (e.g., laugh = happy, cry = sad, scream = fear) produced by 10 actors ($n = 5$ women). Each actor’s vocalization was validated by TD adults, producing a reliability coefficient of .67 or greater.

Music. The participants listened to 24 short instrumental melodies or succession of chords played on a violin, from the Musical Emotional Bursts battery (MEB [55]), a validated set of musical clips designed as a musical counterpart to the MAV vocal stimuli. Each selected clip that conveyed happiness, sadness, or fear had a reliability coefficient of .70 or greater validated with TD adults.

Categorical and dimensional response conditions. A three-tiered response system (a categorical condition and a two-part dimensional condition) was provided following the presentation of each stimulus item within the three blocks. The order of response conditions was counterbalanced across the participants, with half of the participants ($n = 24$) responding with the categorical condition first and the other half with the dimensional condition first.

Categorical condition. In the categorical response condition, each item was presented and then followed by the on-screen instruction, “Which emotion?” with three verbal labels of happy, sad, and fear (the word scared was used when depicting emotions evoked through faces and voices, and scary for emotions evoked through music; see Fig 1). The participants selected their response by clicking on one of the three options using a mouse. Accuracy of emotion recognition was determined by the percentage of correct responses per emotion and also per stimulus type.

Dimensional condition. In the dimensional response condition, the presentation of each item was followed by the on-screen instruction, “How positive or negative is the face (or voice or music)?” The participants provided a valence rating on a 5-point Likert scale (from very negative to very positive). They were also asked “How energetic is the face (or voice or music)?” to which they provided an arousal rating on a 5-point Likert scale (from very low energy to very high energy). In order to minimize the verbal demand of the task and to make

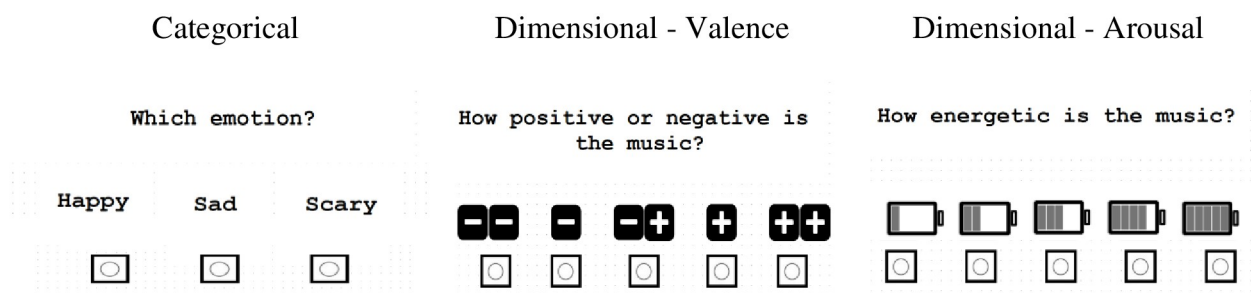


Fig 1. Example categorical and dimensional response options.

<https://doi.org/10.1371/journal.pone.0279002.g001>

the task more child-friendly, dimensional Likert scale anchors were paired with cartoon icons. Addition and subtraction symbols were used for the valence scale and batteries with varying levels of energy were used for the arousal scale (see Fig 1).

Procedure

All of the procedures for this study were approved by McGill University's Research Ethics Board (#101–0716). As Quebec is a bilingual (English and French) province, all of the materials and procedures were available or translated into both languages and administered in the participant's dominant language by bilingual clinically trained research staff. The participants were tested in a research lab at the university or in the child's school or home. Written informed consent was obtained from the parents and written or verbal assent was provided by each participant prior to the beginning of the study.

All of the participants successfully completed a basic hearing test at www.legroupeforget.com. The hearing test and experimental task were administered on a PC laptop with over-ear headphones. The experimental task was completed in the E-prime (version 3.0) environment using a laptop and mouse. In order to ensure the comprehension of both response conditions, a practice session was administered first. All of the instructions were available on the screen in the participants' dominant language and were read to them during the practice session. A more extensive explanation of the Likert scales was provided during the practice session if needed. The experimental task took approximately 20 minutes to complete.

Analysis plan

Separate repeated measures analyses of covariance (ANCOVAs) were run for each of the categorical, dimensional valence, and dimensional arousal conditions, to understand the effects of group (AS, TD), stimulus type (music, face, voice), and emotion (happy, sad, fear), while controlling for VMA. Significant effects were further examined in two ways: 1) between-group comparisons, using two-way repeated measures ANOVAs to examine group differences in task performance for each stimulus type and/or emotion, while controlling for differences in VMA of all the participants; and 2) within-group comparisons, using repeated measures ANOVAs to examine differences in task performance across stimulus types and/or emotions within each group, controlling for VMA for each group. To yield medium-sized effects ($F = .25$) with 80% power and alpha at .05, a priori power analyses using G*Power 3.1 indicated that a total sample size of 46 participants would be needed. Bonferroni corrections for multiple comparisons were applied when appropriate. The task performance means and means adjusted for VMA are presented in Tables 2 and 3.

Results

Categorical response condition

The three-way interaction among group, stimulus type, and emotion on categorical ER accuracy ratings was not statistically significant, $F(2.98, 134.22) = 2.15, p = .098, \eta_p^2 = .045$. A significant main effect of stimulus, $F(1.61, 72.24) = 7.48, p = .002, \eta_p^2 = .14$, and a two-way interaction between group and stimulus, $F(1.61, 72.24) = 6.58, p = .004, \eta_p^2 = .13$, were found when controlling for VMA (see Table 2 and Fig 2). Follow up analyses revealed a significant between-group difference in ER accuracy within the musical stimuli, $F(1,45) = 5.40, p = .025, \eta_p^2 = .11$, as the AS group had higher mean accuracy than the TD group when recognizing emotions presented in music. In contrast, no differences were found between the AS and TD groups on ER accuracy within the vocal stimuli, $F(1, 45) = .22, p = .64, \eta_p^2 = .005$, or the facial

Table 2. Means, adjusted means, standard deviations and standard errors for categorical ER accuracy ratings.

	AS Group		TD Group		F	η_p^2
	Mean (SD)	Adjusted Mean (SE)	Mean (SD)	Adjusted Mean (SE)		
Stimulus Type					6.58*	.13
Music	.78 (.14)	.80 (.03)	.70 (.18)	.69 (.03)	5.40*	.11
Face	.80 (.14)	.81 (.01)	.85 (.09)	.84 (.01)	.94	.02
Voice	.92 (.07)	.92 (.03)	.91 (.06)	.91 (.03)	.22	.005
Emotion					3.05	.06
Happy	.89 (.13)	.90 (.02)	.87 (.09)	.85 (.02)		
Sad	.81 (.14)	.82 (.03)	.87 (.11)	.86 (.03)		
Fear	.80 (.14)	.81 (.03)	.74 (.18)	.73 (.03)		

Note. Means adjusted for VMA of the overall sample = 9.22 years. Statistics for individual emotions are not reported because the emotion by group interaction effect was not significant.

* $p < .05$

<https://doi.org/10.1371/journal.pone.0279002.t002>

stimuli, $F(1, 45) = .94, p = .34, \eta_p^2 = .02$, suggesting that the children in the AS and TD groups performed equally well in terms of ER accuracy from voices and faces.

The two-way interaction between group and stimulus type was also explored by comparing ER accuracy ratings across the three stimulus types (music, face, voice) within each group. The main effect of stimulus type was significant within the TD group, $F(1.31, 27.43) = 6.98, p = .009, \eta_p^2 = .25$, but not within the AS group, $F(2, 46) = 1.09, p = .34, \eta_p^2 = .045$.

Table 3. Means, adjusted means, standard deviations and standard errors for dimensional ratings of valence and arousal.

	AS Group		TD Group		F	η_p^2
	Mean (SD)	Adjusted Mean (SE)	Mean (SD)	Adjusted Mean (SE)		
Valence Ratings						
Stimulus Type					1.55	.03
Music	3.00 (.57)	2.97 (.10)	2.80 (.37)	2.83 (.10)		
Face	2.88 (.35)	2.87 (.07)	2.93 (.31)	2.94 (.07)		
Voice	2.93 (.46)	2.93 (.09)	2.84 (.37)	2.85 (.09)		
Emotion					4.30*	.09
Happy	4.32 (.34)	4.33 (.09)	3.91 (.47)	3.90 (.09)	11.77*	.21
Sad	2.12 (.60)	2.09 (.11)	2.19 (.50)	2.22 (.12)	.57	.01
Fear	2.37 (.75)	2.34 (.14)	2.47 (.54)	2.50 (.14)	.66	.01
Arousal Ratings						
Stimulus Type					2.47	.05
Music	3.20 (.46)	3.26 (.10)	3.15 (.59)	3.08 (.11)		
Face	3.03 (.35)	3.04 (.08)	3.15 (.42)	3.14 (.09)		
Voice	3.20 (.60)	3.23 (.12)	3.22 (.53)	3.18 (.12)		
Emotion					1.87	.04
Happy	4.15 (.48)	4.12 (.10)	3.81 (.50)	3.84 (.11)		
Sad	2.41 (.59)	2.44 (.12)	2.65 (.56)	2.62 (.13)		
Fear	2.87 (.86)	2.97 (.16)	3.04 (.77)	2.94 (.17)		

Means adjusted for VMA of the overall sample = 9.22 years. Statistics for individual emotions and stimulus types are not reported when corresponding group by stimulus type or group by emotion interactions are not significant.

* $p < .05$

<https://doi.org/10.1371/journal.pone.0279002.t003>

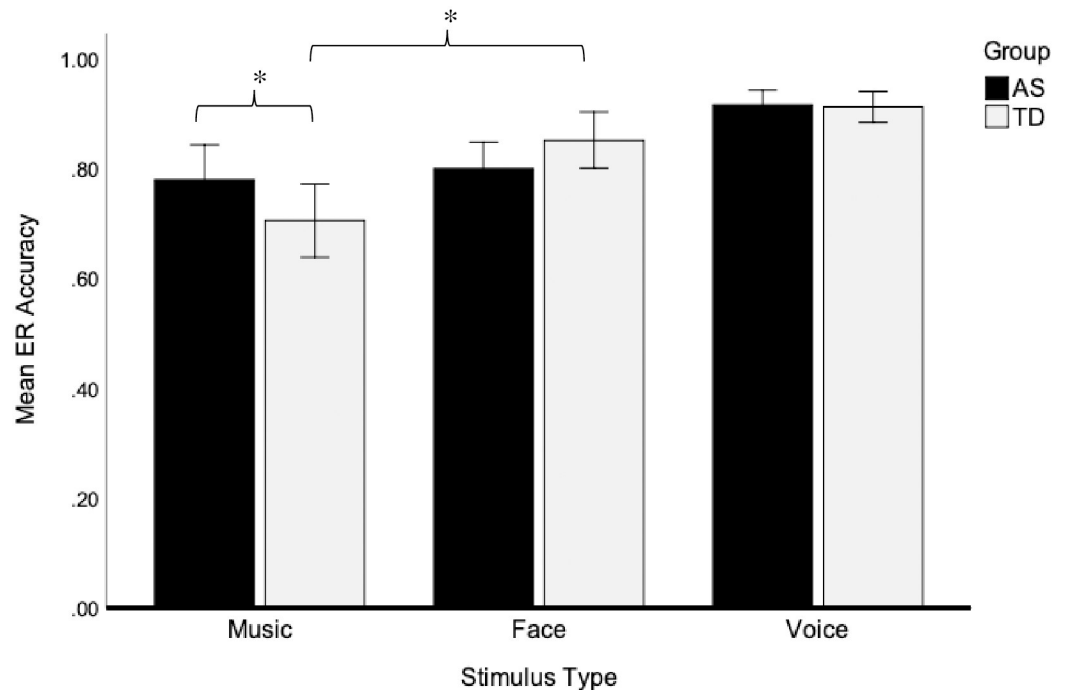


Fig 2. Between- and within-group differences across stimulus types in the categorical condition. Musical emotion recognition (ER) accuracy was significantly higher among the children on the AS as compared to the TD children, whereas the groups did not differ in ER from voices or faces. Although both groups had higher ER accuracy from voices, the TD group had significantly lower ER accuracy from music than faces whereas no differences emerged between music and faces for the AS group. Means and standard error bars are shown. * $p < .05$.

<https://doi.org/10.1371/journal.pone.0279002.g002>

Nevertheless, given our a priori hypothesis, we compared responses across stimulus types within both groups. In both groups, ER accuracy was highest for emotions from voices (AS group: $M = .92$, TD group: $M = .91$), $ps \leq .001$. The children in the TD group recognized emotions more accurately from faces compared to music, $p < .001$, whereas no significant difference was found between ER accuracy from faces compared to music for the AS group, $p = .36$.

Neither the main effects of group and emotion nor the interaction effects of emotion by stimulus type and emotion by group were significant (all p values $> .05$), suggesting that happy, sad, and fear emotions were equally well identified within and between the AS and TD groups, and across the stimulus types. VMA was not a significant covariate in the model, and neither the interaction effects of VMA by emotion nor of VMA by stimulus type were significant (all p values $> .05$).

These results support our hypothesis that children on the AS show a strength in identifying emotions within the musical stimuli compared to TD children. However, contrary to another hypothesis, the TD children did not demonstrate a strength in ER from facial and vocal stimuli as compared to the children on the AS. Our within group hypotheses were also not supported. Specifically, the children on the AS did not demonstrate a relative weakness in facial or vocal ER as compared to musical ER. Both groups had better ER accuracy from voices, and the children on the AS performed comparably between faces and music, whereas the TD children performed better with faces than music. Overall, the children on the AS showed a strength in recognizing emotions in music in comparison to the TD children, but not in comparison to their ability to recognize emotions in faces and voices.

Dimensional response condition

Valence ratings. A three-way interaction among group, stimulus type, and emotion on valence ratings was not statistically significant, $F(4, 180) = .86, p = .49, \eta_p^2 = .02$. Neither a significant main effect of emotion nor a significant two-way interaction between group and emotion, $F(1.52, 68.35) = 4.30, p = .026, \eta_p^2 = .09$, were found while controlling for VMA. Follow up analyses revealed a significant group difference in valence ratings for happy emotions, $F(1, 45) = 11.77, p = .001, \eta_p^2 = .21$, such that the AS group rated happy emotions as significantly more positive as compared to the TD group. In contrast, no differences were found between the AS and TD groups on mean valence ratings of sad, $F(1, 45) = .57, p = .46, \eta_p^2 = .01$, or fear emotions, $F(1, 45) = .66, p = .42, \eta_p^2 = .01$, suggesting they rated emotional valence for these two negative emotions similarly (see Table 3 and Fig 3).

The main effects of group and stimulus type, and the interaction effects of stimulus type by group and stimulus type by emotion were not significant (all p values $>.05$), suggesting that the valence of emotions for both groups was equally well identified regardless of the stimuli in which they were presented. VMA was not a significant covariate in the model, and the interaction effects between VMA and emotion and stimulus type were not significant (all p values $>.05$).

These results suggested that both groups rated the valence of emotions similarly (i.e., happy emotions as more positive than sad and fear), regardless of the stimuli in which they were presented, although the children on the AS rated happy emotions as more positive than did the TD children.

Arousal ratings. A three-way interaction among group, stimulus type, and emotion on arousal ratings was not statistically significant, $F(4, 180) = 1.01, p = .40, \eta_p^2 = .022$. Further, the two-way interactions between group and stimulus type, $F(2, 90) = 2.47, p = .09, \eta_p^2 = .052$, and

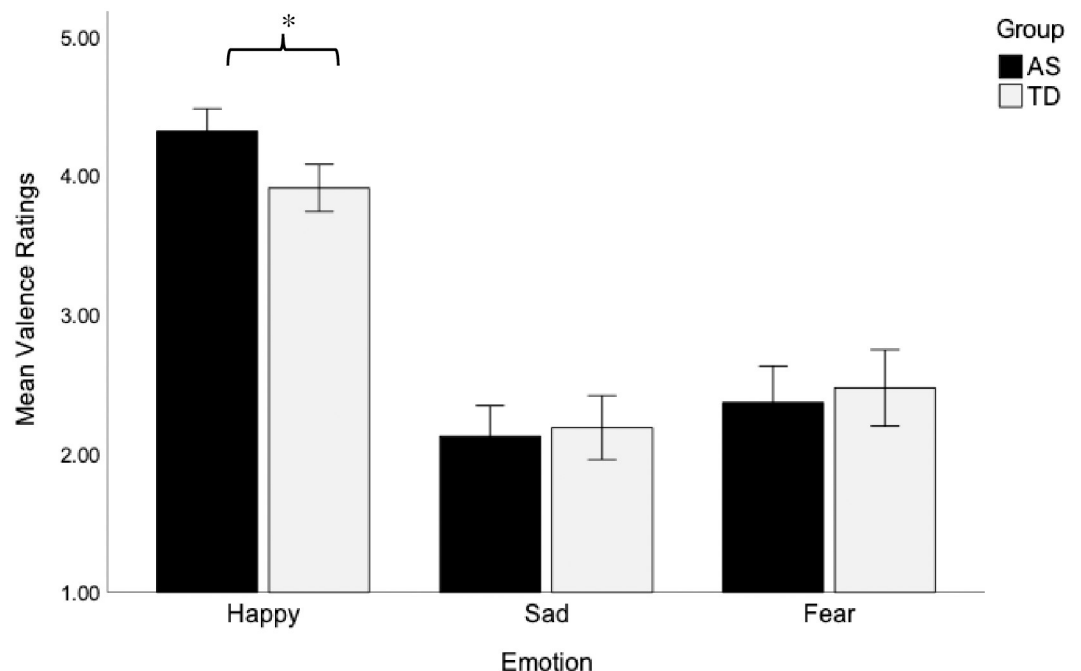


Fig 3. Between-group difference in valence ratings for happy emotions. The children on the AS rated happy stimuli as more positive than did the TD children, whereas the groups did not differ in valence ratings of sad and fear emotions. Means and standard error bars are shown. * $p = .001$.

<https://doi.org/10.1371/journal.pone.0279002.g003>

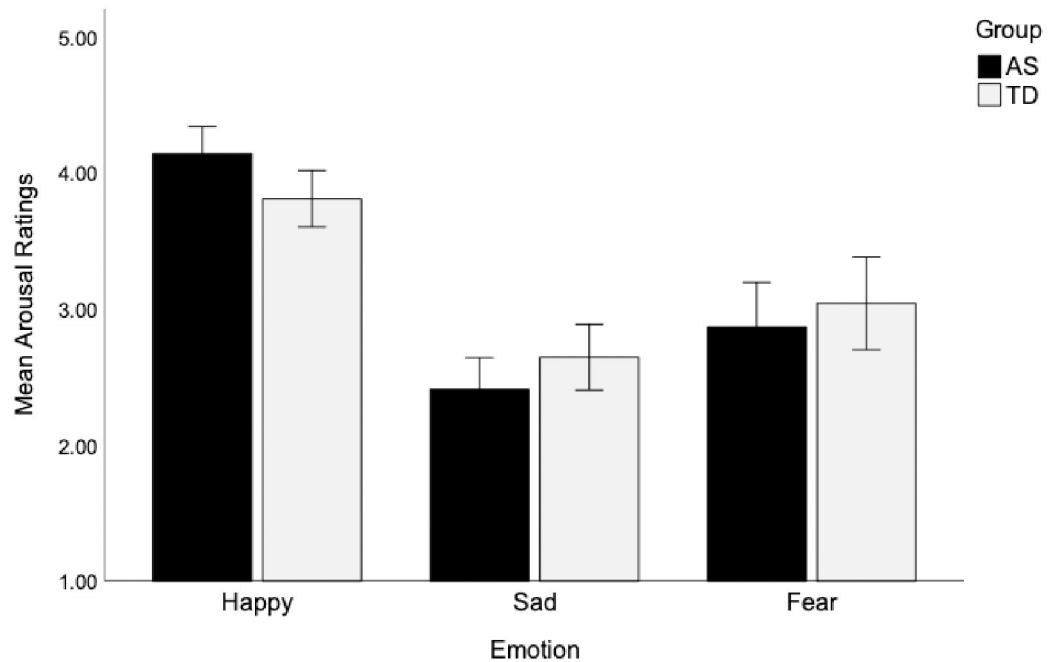


Fig 4. No group differences in arousal ratings for emotions. Both AS and TD groups rated happy emotions as more arousing than fear, and fear more arousing than sad. Means and standard error bars are shown.

<https://doi.org/10.1371/journal.pone.0279002.g004>

group and emotion, $F(2, 90) = 1.87, p = .17, \eta_p^2 = .04$, were not significant. These results suggest that the children on the AS and the TD children showed comparable patterns of arousal ratings across emotions and stimulus types (see Table 2 and Fig 4).

A significant main effect of emotion on arousal ratings was found overall, $F(1.65, 74.22) = 13.97, p < .001, \eta_p^2 = .24$, and within each group separately (AS group: $F(2, 46) = 7.71, p = .001, \eta_p^2 = .25$; TD group: $F(1.58, 33.26) = 6.63, p = .006, \eta_p^2 = .24$). Both of the groups rated happy emotions (AS: $M = 4.12$; TD: $M = 3.84$) as more arousing than fear (AS: $M = 2.97$; TD: $M = 2.94$), and fear as more arousing than sad (AS: $M = 2.65$; TD: $M = 2.62$), all p values $\leq .005$, after controlling for VMA. Neither the main effects of group and stimulus type, nor the interaction effect of stimulus type by emotion were significant (p values $> .05$), suggesting that the arousal ratings of emotions for children in both groups were equally well identified regardless of the stimulus type in which they were presented.

Although VMA was not a significant covariate in the overall model ($p = .14$), significant interactions were found between VMA and emotion, $F(1.65, 74.22) = 4.96, p = .014, \eta_p^2 = .099$, and between VMA and stimulus type, $F(2, 90) = 3.33, p = .04, \eta_p^2 = .069$. One-tailed Pearson correlations revealed that VMA was positively correlated with arousal ratings of fear ($r = .37, p = .005$), such that children with higher VMA rated fearful emotions as more arousing than those with lower VMA (Bonferroni corrected $p = .0167$ to account for multiple comparisons). No other correlations with emotion or stimulus type reached significance. These results suggested that VMA may have contributed to patterns of arousal ratings for particular emotions (i.e., fear).

Discussion

Our primary objective in this study was to reconcile contrasting findings in the literature on emotion recognition (ER) among persons on the AS that include purported challenges or

“deficits” in basic ER with socially explicit stimuli (primarily faces and voices) and relative strengths in ER with musical stimuli. We aimed to determine whether relative ER strengths among children on the AS were specific to music, or whether assumptions of difficulty associated with facial and vocal ER needed re-examination methodologically, via a direct group comparison of performance on a traditional categorical behavioral ER task involving musical, facial, and vocal stimuli. Music afforded a unique opportunity to compare ER skills from music to conventional ER paradigms, as musical stimuli can convey emotions without reliance on the usual interpersonal social context. Additionally, we attempted to enhance our understanding of the abilities and styles of emotion processing among children on the AS by also applying a nontraditional dimensional approach to our multi-stimulus investigation of ER. As such, our secondary objective was to take an exploratory look at patterns of performance between and within the groups using continuous ratings of valence and arousal among music, faces, and voices.

Based on the literature related to our first objective, we offered data-driven hypotheses that children on the AS as compared to TD children would demonstrate better accuracy with musical ER, and reduced accuracy with facial and vocal ER. We also hypothesized that within the groups, the children on the AS would demonstrate a strength with musical ER as compared to facial and vocal ER, whereas we expected that the TD children would perform equally well regardless of stimulus type. Given the exploratory nature of our secondary objective, we did not offer specific hypotheses about patterns of responding using dimensional valence and arousal ratings between the groups. To test our hypotheses, we developed an integrated paradigm to examine ER skills with musical, facial, and vocal stimuli conveying happy, sad, and fear emotions, using both categorical and dimensional response options concurrently.

Patterns of basic ER among children on the AS and TD children

Within the categorical condition, we found that the children on the AS demonstrated greater accuracy in musical ER compared to the TD children, when controlling for verbal mental age (VMA), thereby supporting findings of strengths in musical ER among children on the AS [8–10]. However, in contrast with prior meta-analytic findings [6, 7, 31], the children on the AS and the TD children identified emotions equally well when using the facial and the vocal stimuli. Further, patterns of within-group performance revealed that children in both groups demonstrated the greatest ER accuracy with vocal stimuli, but that the children on the AS demonstrated comparable ER accuracy between faces and music whereas the TD children showed reduced ER accuracy with music as compared to with faces. Within the dimensional condition, the groups also performed comparably, with one exception. In terms of valence ratings, the children on the AS rated happy emotions as more positive than the TD children, but otherwise the groups rated the valence of emotions similarly across the three stimulus types. Further, both groups displayed similar patterns of arousal ratings for the three emotions and across the three stimulus types. VMA did not appear to add any further influence to our findings except for being positively correlated with arousal ratings of fear.

Considering possible explanations for our pattern of findings that stand in contrast with much of the previous ER literature necessitates a relevant discussion of methodological issues, especially as null results were found. Although we found clear discrepancies between the groups on the musical ER task, in which the AS group outperformed the TD group, our tasks may not have been sensitive enough to detect meaningful differences between the groups, thus resulting in comparable ER accuracy patterns for faces and voices. Specifically with regard to our vocal ER task, we used the most basic measure of vocal ER involving short, non-verbal vocalizations (i.e., cries, screams, and laughs) rather than verbal vocalizations which require

processing of linguistic aspects of semantic content (e.g., “I am afraid”) or prosody of speech (e.g., “That’s a big shark!” said in a fearful tone), in order to understand the emotion being conveyed. Although minimizing the verbal component of the task provided confidence that ER from voices was not disproportionately reliant on participants’ verbal skills, it may have contributed to high performance accuracy overall. Nonetheless, our findings of comparable performance for ER in faces and voices are consistent with studies of socially explicit ER that also account for VMA or IQ (e.g., [40, 57–59]).

The number and type of emotion choices provided also contributes to task difficulty. Our study included three basic emotions that can easily be conveyed in music, as well as in faces and voices (and thus offered a 33% chance of randomly guessing the correct answer), with a single positive emotion and two negative emotions. Group differences may be more readily apparent when measuring complex or higher order emotion processing (e.g., [60, 61]). Further, the evidence from studies with more complex emotions and more choices (and therefore decreased accuracy attributable to chance) indicate that TD groups also perform less accurately relative to their performance with basic emotions and fewer choices [62, 63]. However, our integrated categorical and dimensional task could be extended to measure ER responses among more complex facial and vocal stimuli and to survey both basic and complex emotions (e.g., [64, 65]).

Furthermore, by design, the dimensional condition does not lend itself to constrained and categorical performance metrics, but rather allows for examination of general patterns of response using 5-point Likert scales to measure continuous components of valence and arousal. Of the 12 comparisons across two dimensions, three emotions, and three stimulus types, the only group difference was found in the perception of valence for happy emotions. The patterns of continuous valence and arousal ratings across all other emotions and stimulus types were virtually indistinguishable. Nevertheless, a lack of group differences on behavioral tasks does not necessarily exclude underlying neurophysiological differences exhibited during emotion processing. Evidence from studies of neuroimaging, skin conductance, and pupil dilation have highlighted differences in arousal levels during emotion processing tasks between TD persons and persons on the AS who also have Type II alexithymia, a difficulty with appraising and verbalizing emotions without necessarily experiencing such emotions in an atypical manner (see [52] for a review). Thus, extensions of the current study would benefit from the comparison of potential discrepancies between neurophysiological processing of emotions and the associated behavioral ratings.

Alternate ecological and sociocultural perspectives regarding why the performance of children in our study was highly accurate overall should also be considered. For example, the impact of verbal cognitive ability on ER is being increasingly considered and accounted for in research comparing children with and without developmental disabilities. Further, access to educational curricula and school-based interventions that prioritize social emotional learning (SEL), likely play a significant role in overall improvements in verbally mediated emotion learning and communication. In the United States, for example, such widespread proliferation of hundreds of universal school-based K-12 SEL programs—supported by federal educational policies—has occurred over the past few decades. Such SEL programs have led to demonstrated improvements in social and emotional skills, as well as attitudes, behaviors, and academic performance among students with a diversity of processing styles and needs [66], including those with mental health challenges such as anxiety and depression, or those with difficulty verbally expressing emotions, such as in alexithymia. Similarly, the proliferation in the past few decades of evidence-based emotion-focused interventions specifically for children on the AS that highlight the importance of identifying, expressing, and regulating emotions (e.g., PEERS [67]; SCERTS Model [68]; Facing Your Fears [69]; EASE [70]) further demonstrate increasing

societal priorities surrounding the importance of research, policy, and funding for social and emotional learning for all children.

Neuroconstructivist approaches and the case for equifinality in ER

Rather than demonstrating a singular mechanism and pattern of processing by which both groups achieve comparable task performance, our findings may instead highlight unique styles of processing for each group that lead to similar outcomes. Cicchetti and Rogosch [71] argued that the principles of multifinality and equifinality allow us to better understand possible mechanisms for observable outcomes among different populations. Whereas multifinality infers that similar developmental circumstances may lead to a heterogeneity of outcomes, equifinality suggests that various circumstances or abilities may lead to similar outcomes. The lens of equifinality specifically has been readily applied to conceptualize the heterogeneity of both genetic and behavioral characteristics of persons on the AS (i.e., multiple factors may predispose an individual to be on the AS). For example, Burack et al. [72] and others have argued that persons on the AS may engage in different styles of cognitive processing than TD persons, and thus call for greater exploration of the various ways persons on the AS engage with stimuli, rather than viewing cognitive or task performance as “deficient” or atypical compared to a reference standard. Relatedly, Johnson et al. [73, 74] have posited that the behavioral characteristics of persons on the AS may manifest as adaptive responses to early differences in neural, genetic, and experience-dependent development, and therefore impact subsequent developmental trajectories and styles of information processing. For example, differences in brain region growth or connectivity may prompt information processing reorganization (e.g., [75, 76]) in a manner that is adaptive and optimized for relevant or preferred experience-dependent learning.

The usefulness of considering neuroconstructivist approaches and applying a lens of equifinality to understand differing mechanisms of cognitive processing and ER among persons on the AS and TD persons can be highlighted in specific conceptual approaches, such as the Enhanced Perceptual Functioning (EPF) model [77, 78]. According to the EPF model, persons on the AS may have a default or preferential bottom-up approach to processing the perceptual features of a stimulus that favors local over global aspects. In this framework, global processing among persons on the AS is not necessarily impaired but may be deprioritized relative to local processing. In comparison, TD persons may favor a top-down global or holistic processing style, and may deprioritize processing at the level of local perceptual cues. Such processing differences may be an optimized or adaptive response to early developmental differences in neural structure or connectivity [79] and may be further reinforced by continued exposure and engagement with salient and motivating cues. For example, TD children spend more time looking at socially explicit stimuli (e.g., caregiver interactions, videos of social scenes) than do children on the AS [80] and therefore gain greater experience with emotional content conveyed in faces and voices. By comparison, children on the AS tend to demonstrate enhanced local processing of perceptual features of socially explicit faces [81, 82], speech [83] and audiovisual synchronous biological motion [84], in addition to their enhanced perceptual processing of less explicitly social stimuli, such as music [79]. Further, processing of emotional information through preferred or engaging stimuli such as music, which activates reward-based regions of the brain [13, 29], may thus become more salient and reinforcing—and thus, further optimized—over time.

Across all three stimulus types, basic emotions can be decoded using both top down and bottom up approaches, and thus both groups of children in our study may have been successful at decoding emotions by using their preferential styles of processing in order to categorize emotions similarly on our task. Bottom-up processing of music may confer an additional

advantage to children on the AS, who show strengths in decoding perceptual musical cues (i.e., pitch, tone, energy) (e.g., [11, 12]). Thus, compared to TD children, music may readily invoke the enhanced perceptual processing styles of children on the AS and be a privileged means for conveying emotions for them.

Implications

For a characteristic or challenge to be considered a core “deficit” primary to any particular group, it should be specific to and universal within that group, and should be one of the most persistent or impairing features [40, 85]. Overall, our findings are inconsistent with the notion that children on the AS demonstrate a core “deficit” in basic ER, and we attempt to delineate this through our findings in several ways. In terms of specificity, using a child-friendly, multi-response paradigm designed to minimize reliance on verbal cues, we found typical ER performance across the facial and vocal accuracy tasks, and comparable ratings of emotional valence and arousal, demonstrating indisputable basic ER capabilities of children on the AS. The notion of the universality of ER differences was also not borne out, given our findings of relative strength of children on the AS in musical ER, which ultimately underscore that discussions of ER broadly need not be restricted to the traditional facial or vocally expressed emotions.

The musical ER strengths highlighted in our study support the growing evidence of the benefits of music therapy for children on the AS [27]. Our findings contribute to the notion that musical strengths, including typical or enhanced processing of music-evoked emotions, may be a key component of the success of music programs and therapies that target social communication and interaction skills [23, 26–28] and speech and language interventions that incorporate music [22, 24, 25]. Music therapy, education, and interventions may be successful because they leverage emotional and cognitive strengths [13] and are motivating and accessible to children on the AS because they readily engage with and enjoy music [86, 87].

Conclusion

The findings from this study of ER across multiple basic emotions, stimulus types, and response options, help to dispel notions of deficits of the ability to explicitly recognize and label basic emotions of children on the AS relative to TD children. Instead, findings of comparable ER accuracy (in faces and voices) or relative strengths (in musical ER) depending on the stimulus type provides support for the use of music therapy, education, and interventions. Ultimately, our findings contribute to efforts to provide a nuanced understanding of the ways in which we can re-shape our thinking about persons on the AS by engaging their preferences and strengths.

Supporting information

S1 Table. Family demographics for the AS and TD groups. Family demographics data were available for 36 of 48 participants. **p* values of Fisher’s exact tests between the AS and TD groups were not significant following Bonferroni corrections for multiple comparisons. (DOCX)

Acknowledgments

We are thankful to the graduate students and research assistants in the Behaviour, Autism and NeuroDevelopment (BAND) Research Group who assisted with data collection, and to the children and families who participated in our study.

Author Contributions

Conceptualization: Shalini Sivathasan, Eve-Marie Quintin.

Formal analysis: Shalini Sivathasan, Hadas Dahary.

Funding acquisition: Shalini Sivathasan, Eve-Marie Quintin.

Methodology: Shalini Sivathasan.

Project administration: Eve-Marie Quintin.

Resources: Eve-Marie Quintin.

Software: Eve-Marie Quintin.

Supervision: Jacob A. Burack, Eve-Marie Quintin.

Writing – original draft: Shalini Sivathasan.

Writing – review & editing: Shalini Sivathasan, Hadas Dahary, Jacob A. Burack, Eve-Marie Quintin.

References

1. Baron-Cohen S. (1988). Social and pragmatic deficits in autism: Cognitive or affective? *Journal of Autism and Developmental Disorders*, 18(3), 379. <https://doi.org/10.1007/BF02212194> PMID: 3049519
2. Baron-Cohen S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences*, 6(6), 248–254. [https://doi.org/10.1016/s1364-6613\(02\)01904-6](https://doi.org/10.1016/s1364-6613(02)01904-6) PMID: 12039606
3. Baron-Cohen S., Ring H., Wheelwright S., Bullmore E. T., Brammer M.J., Simmons A., et al. (1999). Social intelligence in the normal and autistic brain: an fMRI study. *European Journal of Neuroscience*, 11, 1891–98. <https://doi.org/10.1046/j.1460-9568.1999.00621.x> PMID: 10336657
4. Hobson R. P. (1986a). The autistic child's appraisal of expressions of emotion. *Journal of Child Psychology and Psychiatry*, 27(3) 321–342. <https://doi.org/10.1111/j.1469-7610.1986.tb01836.x> PMID: 3733915
5. Hobson R. P. (1986b). The autistic child's appraisal of expressions of emotion: A further study. *Journal of Child Psychology and Psychiatry*, 27(5), 671–680. <https://doi.org/10.1111/j.1469-7610.1986.tb00191.x> PMID: 3771683
6. Lozier L. M., Vanmeter J. W., & Marsh A. A. (2014). Impairments in facial affect recognition associated with autism spectrum disorders: a meta-analysis. *Development and Psychopathology*, 26(4pt1), 933–945. <https://doi.org/10.1017/S0954579414000479> PMID: 24915526
7. Uljarevic M., & Hamilton A. (2012). 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> PMID: 23114566
8. Heaton P., Hermelin B., & Pring L. (1999). Can children with autistic spectrum disorder perceive affect in music? An experimental investigation. *Psychological Medicine*, 29, 1405–1410. <https://doi.org/10.1017/S0033291799001221> PMID: 10616946
9. Quintin E. M., Bhatara A., Poissant H., Fombonne E., & Levitin D. J. (2011). Emotion perception in music in high-functioning adolescents with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 41(1240–1255). <https://doi.org/10.1007/s10803-010-1146-0> PMID: 21181251
10. Stephenson K. G., Quintin E. M., & South M. (2016). Age-related differences in response to music-evoked emotion among children and adolescents with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 46(4), 1142–1151. <https://doi.org/10.1007/s10803-015-2624-1> PMID: 26520146
11. Heaton P. (2008). Assessing musical skills in autistic children who are not savants. *Philosophical Transactions of the Royal Society: Biological Sciences*, 364(1522), 1443–1447. <https://doi.org/10.1098/rstb.2008.0327> PMID: 19528029
12. Mottron L., Bouvet L., Bonnel A., Samson F., Burack J. A., Dawson M., et al. (2013). Veridical mapping in the development of exceptional autistic abilities. *Neuroscience & Biobehavioral Reviews*, 37(2), 209–228. <https://doi.org/10.1016/j.neubiorev.2012.11.016> PMID: 23219745

13. Quintin E. M. (2019). Music-evoked reward and emotion: Relative strengths and response to intervention of people with ASD. *Frontiers in Neural Circuits*, 13, 49. <https://doi.org/10.3389/fncir.2019.00049> PMID: 31619969
14. Bonnel A., Mottron L., Peretz I., Trudel M., Gallun E., & Bonnel A. M. (2003). Enhanced pitch sensitivity in individuals with autism: a signal detection analysis. *Journal of Cognitive Neuroscience*, 15(2), 226–235. <https://doi.org/10.1162/089892903321208169> PMID: 12676060
15. Bonnel A., McAdams S., Smith B., Berthiaume C., Bertone A., Ciocca V., et al. (2010). Enhanced pure-tone pitch discrimination among persons with autism but not Asperger syndrome. *Neuropsychologia*, 48(9), 2465–2475. <https://doi.org/10.1016/j.neuropsychologia.2010.04.020> PMID: 20433857
16. Heaton P. (2003). Pitch memory, labelling and disembedding in autism. *Journal of Child Psychology and Psychiatry*, 44(4), 543–551. <https://doi.org/10.1111/1469-7610.00143> PMID: 12751846
17. Heaton P. (2005). Interval and contour processing in autism. *Journal of Autism and Developmental Disorders*, 35(6), 787. <https://doi.org/10.1007/s10803-005-0024-7> PMID: 16283085
18. Mottron L., Peretz I., & Menard E. (2000). Local and global processing of music in high-functioning persons with autism: beyond central coherence? *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 41(8), 1057–1065. <https://doi.org/10.1111/1469-7610.00693> PMID: 11099122
19. Heaton P., Hermelin B., & Pring L. (1998). Autism and pitch processing: A precursor for savant musical ability? *Music Perception*, 15(3), 291–305. <https://doi.org/10.2307/40285769>
20. Heaton P., Williams K., Cummins O. & Happe F. (2008). Autism and pitch processing splinter skills: A group and sub-group analysis. *Autism*, 12, 21–37.
21. Stanutz S., Wapnick J., & Burack J. A. (2014). Pitch discrimination and melodic memory in children with autism spectrum disorders. *Autism*, 18(2), 137–147. <https://doi.org/10.1177/1362361312462905> PMID: 23150888
22. Chenausky K., Norton A., Tager-Flusberg H., & Schlaug G. (2016). Auditory-motor mapping training: Comparing the effects of a novel speech treatment to a control treatment for minimally verbal children with autism. *PLoS One*, 11(11), e0164930. <https://doi.org/10.1371/journal.pone.0164930> PMID: 27829034
23. Lense M. D., & Camarata S. (2020). PRESS-play: musical engagement as a motivating platform for social interaction and social play in young children with ASD. *Music & Science*, 3. <https://doi.org/10.1177/2059204320933080> PMID: 32832103
24. Lim H. A. (2010). Effect of “developmental speech and language training through music” on speech production in children with autism spectrum disorders. *Journal of Music Therapy*, 47(1), 2–26. <https://doi.org/10.1093/jmt/47.1.2> PMID: 20635521
25. Lim H. A., & Draper E. (2011). The effects of music therapy incorporated with applied behavior analysis verbal behavior approach for children with autism spectrum disorders. *Journal of Music Therapy*, 48(4), 532–550. <https://doi.org/10.1093/jmt/48.4.532> PMID: 22506303
26. Kim J., Wigram T., & Gold C. (2008). The effects of improvisational music therapy on joint attention behaviors in autistic children: a randomized controlled study. *Journal of Autism and Developmental Disorders*, 38(9), 1758–1766. <https://doi.org/10.1007/s10803-008-0566-6> PMID: 18592368
27. LaGasse A. B. (2017). Social outcomes in children with autism spectrum disorder: A review of music therapy outcomes. *Patient Related Outcome Measures*, 8, 23. <https://doi.org/10.2147/PROM.S106267> PMID: 28260959
28. Sharda M., Tuerk C., Chowdhury R., Jamey K., Foster N., Custo-Blanch M., et al. (2018). Music improves social communication and auditory–motor connectivity in children with autism. *Translational Psychiatry*, 8(1), 1–13. <https://doi.org/10.1038/s41398-018-0287-3> PMID: 30352997
29. Caria A., Venuti P., & de Falco S. (2011). Functional and dysfunctional brain circuits underlying emotional processing of music in autism spectrum disorders. *Cerebral Cortex*, 21(12), 2838–2849. <https://doi.org/10.1093/cercor/bhr084> PMID: 21527791
30. Molnar-Szakacs I., & Heaton P. (2012). Music: A unique window into the world of autism. *Annals of the New York Academy of Sciences*, 1252(1), 318–324. <https://doi.org/10.1111/j.1749-6632.2012.06465.x> PMID: 22524374
31. Velikonja T., Fett A. K., & Velthorst E. (2019). Patterns of nonsocial and social cognitive functioning in adults with autism spectrum disorder: A systematic review and meta-analysis. *JAMA Psychiatry*, 76(2), 135–151. <https://doi.org/10.1001/jamapsychiatry.2018.3645> PMID: 30601878
32. Jones W., Carr K., & Klin A. (2008). Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. *Archives of General Psychiatry*, 65(8), 946–954. <https://doi.org/10.1001/archpsyc.65.8.946> PMID: 18678799
33. Klin A., Jones W., Schultz R., & Volkmar F. (2003). The enactive mind, or from actions to cognition: lessons from autism. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1430), 345–360. <https://doi.org/10.1098/rstb.2002.1202> PMID: 12639332

34. Pelphrey K. A., Sasson N. J., Reznick J. S., Paul G., Goldman B. D., & Piven J. (2002). Visual scanning of faces in autism. *Journal of Autism and Developmental Disorders*, 32(4), 249–261. <https://doi.org/10.1023/a:1016374617369> PMID: 12199131
35. Griffin C., Lombardo M. V., & Auyeung B. (2016). Alexithymia in children with and without autism spectrum disorders. *Autism Research*, 9(7), 773–780. <https://doi.org/10.1002/aur.1569> PMID: 26426084
36. Swaminathan S., & Schellenberg E. G. (2015). Current emotion research in music psychology. *Emotion Review*, 7(2), 189–197. <https://doi.org/10.1177/1754073914558282>
37. Harms M. B., Martin A., & Wallace G. L. (2010). Facial emotion recognition in autism spectrum disorders: A review of behavioral and neuroimaging studies. *Neuropsychology Review*, 20(3), 290–322. <https://doi.org/10.1007/s11065-010-9138-6> PMID: 20809200
38. Jones C. R., Pickles A., Falcaro M., Marsden A. J., Happé F., Scott S. K., et al. (2010). A multimodal approach to emotion recognition ability in autism spectrum disorders. *Journal of Child Psychology and Psychiatry*, 52(3), 275–285. <https://doi.org/10.1111/j.1469-7610.2010.02328.x> PMID: 20955187
39. Nuske H. J., Vivanti G., & Dissanayake C. (2013). Are emotion impairments unique to, universal, or specific in autism spectrum disorder? A comprehensive review. *Cognition & Emotion*, 27(6), 1042–1061. <https://doi.org/10.1080/02699931.2012.762900> PMID: 23387530
40. Ozonoff S., Pennington B. F., & Rogers S. J. (1990). Are there emotion perception deficits in young autistic children? *Journal of Child Psychology and Psychiatry*, 31(3), 343–361. <https://doi.org/10.1111/j.1469-7610.1990.tb01574.x> PMID: 2318918
41. Ekman P., & Friesen W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, 1(1), 49–98. <https://doi.org/10.1515/semi.1969.1.1.49>
42. Ekman P., & Friesen W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17(2), 124. <https://doi.org/10.1037/h0030377> PMID: 5542557
43. Russell J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39, 1161–1178. <https://doi.org/10.1037/h0077714>
44. Posner J., Russell J. A., & Peterson B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, 17(03). <https://doi.org/10.1017/S0954579405050340> PMID: 16262989
45. Tseng A., Bansal R., Liu J., Gerber A. J., Goh S., Posner J., et al. (2014). Using the circumplex model of affect to study valence and arousal ratings of emotional faces by children and adults with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 44(6), 1332–1346. <https://doi.org/10.1007/s10803-013-1993-6> PMID: 24234677
46. Fombonne E., Zakarian R., Bennett A., Meng L., & Mclean-Heywood D. (2006). Pervasive developmental disorders in Montreal, Quebec, Canada: Prevalence and links with immunizations. *Pediatrics*, 118(1). <https://doi.org/10.1542/peds.2005-2993> PMID: 16818529
47. Lazoff T., Zhong L., Piperni T., & Fombonne E. (2010). Prevalence of pervasive developmental disorders among children at the English Montreal School Board. *The Canadian Journal of Psychiatry*, 55(11), 715–720. <https://doi.org/10.1177/070674371005501105> PMID: 21070699
48. Constantino J. N., & Gruber C. P. (2012). *Social Responsiveness Scale—Second Edition (SRS-2)*. Torrance, CA: Western Psychological Services.
49. Wechsler, D. (2014). *WISC-V: Administration and Scoring Manual*. NCS Pearson, Incorporated.
50. Wechsler D. (2011). *Wechsler Abbreviated Scale of Intelligence—Second Edition (WASI-II)*. San Antonio, TX: NCS Pearson.
51. Russo N., Kaplan-Kahn E. A., Wilson J., Criss A., & Burack J. A. (2021). Choices, challenges, and constraints: A pragmatic examination of the limits of mental age matching in empirical research. *Development and Psychopathology*, 33(2), 727–738. <https://doi.org/10.1017/S0954579420001480> PMID: 33955342
52. Sivathasan S., Fernandes T. P., Burack J. A., & Quintin E. M. (2020). Emotion processing and autism spectrum disorder: A review of the relative contributions of alexithymia and verbal IQ. *Research in Autism Spectrum Disorders*, 77, 101608. <https://doi.org/10.1016/j.rasd.2020.101608>
53. Trevisan D. A., & Birmingham E. (2016). Are emotion recognition abilities related to everyday social functioning in ASD? A meta-analysis. *Research in Autism Spectrum Disorders*, 32, 24–42. <https://doi.org/10.1016/j.rasd.2016.08.004>
54. Belin P., Fillion-Bilodeau S., & Gosselin F. (2008). The Montreal Affective Voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior Research Methods*, 40(2), 531–539. <https://doi.org/10.3758/brm.40.2.531> PMID: 18522064
55. Paquette S., Peretz I., & Belin P. (2013). The “Musical Emotional Bursts”: A validated set of musical affect bursts to investigate auditory affective processing. *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00509> PMID: 23964255

56. Tottenham N., Tanaka J. W., Leon A. C., McCarry T., Nurse M., Hare T. A., et al. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, 168(3), 242–249. <https://doi.org/10.1016/j.psychres.2008.05.006> PMID: 19564050
57. Braverman M., Fein D., Lucci D., & Waterhouse L. (1989). Affect comprehension in children with pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 19(2), 301–316. <https://doi.org/10.1007/BF02211848> PMID: 2745394
58. Buitelaar J., Van Der Wees M., Swaab-Barneveld H., & Jan van der Gaag R. (1999). Verbal memory and performance IQ predict theory of mind and emotion recognition ability in children with autistic spectrum disorders and in psychiatric control children. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 40, 869–881. <https://doi.org/10.1111/1469-7610.00505> PMID: 10509882
59. Castelli F. (2005). Understanding emotions from standardized facial expressions in autism and normal development. *Autism*, 9(4), 428–449. <https://doi.org/10.1177/1362361305056082> PMID: 16155058
60. Golan O., Baron-Cohen S., Hill J. J., & Rutherford M. (2007). The 'Reading the Mind in the Voice' test-revised: A study of complex emotion recognition in adults with and without autism spectrum conditions. *Journal of Autism and Developmental Disorders*, 37(6), 1096–1106. <https://doi.org/10.1007/s10803-006-0252-5> PMID: 17072749
61. Golan O., Baron-Cohen S., & Golan Y. (2008). The 'reading the mind in films' task [child version]: Complex emotion and mental state recognition in children with and without autism spectrum conditions. *Journal of Autism and Developmental Disorders*, 38(8), 1534–1541. <https://doi.org/10.1007/s10803-007-0533-7> PMID: 18311514
62. Efenbein H. A., & Ambady N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, 128, 203–235. <https://doi.org/10.1037/0033-2909.128.2.203> PMID: 11931516
63. Tracy J. L., Robins R. W., Schriber R. A., & Solomon M. (2011). Is emotion recognition impaired in individuals with autism spectrum disorders? *Journal of Autism and Developmental Disorders*, 41(1), 102–109. <https://doi.org/10.1007/s10803-010-1030-y> PMID: 20464465
64. Heaton P., Allen R., Williams K., Cummins O., & Happé F. (2008). Do social and cognitive deficits curtail musical understanding? Evidence from autism and Down syndrome. *British Journal of Developmental Psychology*, 26(2), 171–182. <https://doi.org/10.1348/026151007X206776>
65. Williams D., & Happé F. (2010). Recognising 'social' and 'non-social' emotions in self and others: a study of autism. *Autism*, 14(4), 285–304. <https://doi.org/10.1177/1362361309344849> PMID: 20392782
66. Durlak J. A., Weissberg R. P., Dymnicki A. B., Taylor R. D., & Schellinger K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82(1), 405–432. <https://doi.org/10.1111/j.1467-8624.2010.01564.x> PMID: 21291449
67. Laugeson E. A., Frankel F., Gantman A., Dillon A. R., & Mogil C. (2012). Evidence-based social skills training for adolescents with autism spectrum disorders: The UCLA PEERS program. *Journal of Autism and Developmental Disorders*, 42(6), 1025–1036. <https://doi.org/10.1007/s10803-011-1339-1> PMID: 21858588
68. Prizant B. M., Wetherby A. M., Rubin E., & Laurent A. C. (2003). The SCERTS model: A transactional, family-centered approach to enhancing communication and socioemotional abilities of children with autism spectrum disorder. *Infants & Young Children*, 16(4), 296–316.
69. Reaven J., Blakeley-Smith A., Leuthe E., Moody E., & Hepburn S. (2012). Facing your fears in adolescence: Cognitive-behavioral therapy for high-functioning autism spectrum disorders and anxiety. *Autism Research and Treatment*. <https://doi.org/10.1155/2012/423905> PMID: 23091719
70. Conner C. M., White S. W., Beck K. B., Golt J., Smith I. C., & Mazefsky C. A. (2019). Improving emotion regulation ability in autism: The Emotional Awareness and Skills Enhancement (EASE) program. *Autism*, 23(5), 1273–1287. <https://doi.org/10.1177/1362361318810709> PMID: 30400749
71. Cicchetti D., & Rogosch F. A. (1996). Equifinality and multifinality in developmental psychopathology. *Development and Psychopathology*, 8(4), 597–600. <https://doi.org/10.1017/S0954579400007318>
72. Burack J. A., Russo N., Kovshoff H., Fernandes T. P., Ringo J., Landry O., et al. (2016). How I attend—not how well do I attend: Rethinking developmental frameworks of attention and cognition in autism spectrum disorder and typical development. *Journal of Cognition and Development*, 17, 553–567. <https://doi.org/10.1080/15248372.2016.1197226>
73. Johnson M. H., Jones E. J., & Gliga T. (2015). Brain adaptation and alternative developmental trajectories. *Development and Psychopathology*, 27(2), 425–442. <https://doi.org/10.1017/S0954579415000073> PMID: 25997763
74. Johnson M. H. (2017). Autism as an adaptive common variant pathway for human brain development. *Developmental Cognitive Neuroscience*, 25, 5–11. <https://doi.org/10.1016/j.dcn.2017.02.004> PMID: 28233663

75. Belmonte M. K., Cook E. H., Anderson G. M., Rubenstein J. L., Greenough W. T., Beckel-Mitchener A., et al. (2004). Autism as a disorder of neural information processing: Directions for research and targets for therapy. *Molecular Psychiatry*, 9(7), 646–663. <https://doi.org/10.1038/sj.mp.4001499> PMID: 15037868
76. Lewis J. D., & Elman J. L. (2008). Growth-related neural reorganization and the autism phenotype: a test of the hypothesis that altered brain growth leads to altered connectivity. *Developmental Science*, 11(1), 135–155. <https://doi.org/10.1111/j.1467-7687.2007.00634.x> PMID: 18171375
77. Mottron L., & Burack J. A. (2001). Enhanced perceptual functioning in the development of autism. In Burack J. A., Charman T., Yirmiya N., & Zelazo P. R. (Eds.), *The Development of Autism: Perspectives from Theory and Research* (pp. 131–148). Lawrence Erlbaum Associates Publishers.
78. Mottron L., Dawson M., Soulières I., Hubert B., & Burack J. (2006). Enhanced perceptual functioning in autism: An update, and eight principles of autistic perception. *Journal of Autism and Developmental Disorders*, 36(1), 27–43. <https://doi.org/10.1007/s10803-005-0040-7> PMID: 16453071
79. Ouimet T., Foster N. E., Tryfon A., & Hyde K. L. (2012). Auditory-musical processing in autism spectrum disorders: A review of behavioral and brain imaging studies. *Annals of the New York Academy of Sciences*, 1252(1), 325–331. <https://doi.org/10.1111/j.1749-6632.2012.06453.x> PMID: 22524375
80. Constantino J. N., Kennon-McGill S., Weichselbaum C., Marrus N., Haider A., Glowinski A. L., et al. (2017). Infant viewing of social scenes is under genetic control and is atypical in autism. *Nature*, 547(7663), 340–344. <https://doi.org/10.1038/nature22999> PMID: 28700580
81. Hubl D., Bölte S., Feineis-Matthews S., Lanfermann H., Federspiel A., Strik W., et al. (2003). Functional imbalance of visual pathways indicates alternative face processing strategies in autism. *Neurology*, 61(9), 1232–1237 <https://doi.org/10.1212/01.wnl.0000091862.22033.1a> PMID: 14610126
82. Samson F., Mottron L., Soulières I., & Zeffiro T. A. (2012). Enhanced visual functioning in autism: An ALE meta-analysis. *Human Brain Mapping*, 33(7), 1553–1581. <https://doi.org/10.1002/hbm.21307> PMID: 21465627
83. Järvinen-Pasley A., Wallace G. L., Ramus F., Happé F., & Heaton P. (2008). Enhanced perceptual processing of speech in autism. *Developmental Science*, 11(1), 109–121. <https://doi.org/10.1111/j.1467-7687.2007.00644.x> PMID: 18171373
84. Klin A., Lin D. J., Gorrindo P., Ramsay G., & Jones W. (2009). Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature*, 459(7244), 257–261. <https://doi.org/10.1038/nature07868> PMID: 19329996
85. Zelazo P. D., Burack J. A., Benedetto E., & Frye D. (1996). Theory of mind and rule use in individuals with Down syndrome: A test of the uniqueness and specificity claims. *Journal of Child Psychology and Psychiatry*, 37, 479–484. <https://doi.org/10.1111/j.1469-7610.1996.tb01429.x> PMID: 8735448
86. Allen R., Hill E., & Heaton P. (2009). Hath charms to soothe...? An exploratory study of how high-functioning adults with ASD experience music. *Autism*, 13(1), 21–41. <https://doi.org/10.1177/1362361307098511> PMID: 19176575
87. Bhatara A., Quintin E. M., Fombonne E., & Levitin D. J. (2013). Early sensitivity to sound and musical preferences and enjoyment in adolescents with autism spectrum disorders. *Psychomusicology: Music, Mind, and Brain*, 23(2), 100. <https://doi.org/10.1037/a0033754>