

Adaptation to vocal expressions and phonemes is intact in autism spectrum disorder

Bestelmeyer, Patricia; Williams, Bethan; Lawton, Jennifer; Stefanou, Maria-Elena; Koldewyn, Kami; Klein, Christoph; Biscaldi, Monica

Clinical Psychological Science

DOI:

[10.1177/2167702617748401](https://doi.org/10.1177/2167702617748401)

Published: 01/05/2018

Peer reviewed version

[Cyswllt i'r cyhoeddiad / Link to publication](#)

Dyfyniad o'r fersiwn a gyhoeddwyd / Citation for published version (APA):

Bestelmeyer, P., Williams, B., Lawton, J., Stefanou, M.-E., Koldewyn, K., Klein, C., & Biscaldi, M. (2018). Adaptation to vocal expressions and phonemes is intact in autism spectrum disorder. *Clinical Psychological Science*, 6(3), 372-381. <https://doi.org/10.1177/2167702617748401>

Hawliau Cyffredinol / General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Adaptation to vocal expressions and phonemes is intact in autism spectrum disorder

Patricia E.G. Bestelmeyer^{1*}, Bethan Williams¹, Jennifer J. Lawton¹, Maria-Elena Stefanou², Kami Koldewyn¹, Christoph Klein^{2,3}, Monica Biscaldi²

¹School of Psychology, Bangor University, Bangor, Gwynedd, UK.

²Department of Child and Adolescent Psychiatry, Psychotherapy, and Psychosomatics, Medical Faculty, University of Freiburg, Freiburg, Germany.

³Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, Medical Faculty, University of Cologne, Cologne, Germany.

*Correspondence:

Dr Patricia E.G. Bestelmeyer

School of Psychology, Brigantia

Bangor University, Gwynedd, LL57 2AS, UK

E-mail: p.bestelmeyer@bangor.ac.uk

Phone: +44 (0)1248 383488

Abstract

Several recent studies have demonstrated reduced visual aftereffects, particularly to social stimuli, in autism spectrum disorder (ASD). This putative impairment of the adaptive mechanism in ASD has been put forward as a possible explanation for some of the core social problems children with ASD experience (e.g. facial emotion or identity recognition). We addressed this claim in children with ASD and typically developing children by using an established methodology and morphed auditory stimulus set for eliciting robust aftereffects to vocal expressions and phonemes. While children with ASD were significantly worse at categorising the vocal expressions compared to the control stimuli (phoneme categorisation), aftereffect sizes in both tasks were identical in the two participant groups. Our finding suggests that the adaptation mechanism is not universally impaired in ASD and is therefore not an explanation for the social perception difficulties in ASD.

Keywords: autism; adaptation; aftereffects; voice; emotion

Introduction

Autism spectrum disorder (ASD) is a heritable, neurodevelopmental condition, which is in part characterised by lifelong impairments in social communication and interaction (American Psychiatric Association, 2013) as well as altered sensory processing (Marco, Hinkley, Hill & Nagarajan, 2012). Difficulties with processing the social information from faces and voices are a commonly cited aspect of ASD, particularly with regards to identifying and discriminating between facial identities (e.g. Weigelt, Koldewyn, & Kanwisher, 2012 for a review) and emotional expressions from both faces (e.g. Harms, Martin, & Wallace, 2010 for review; Uljarevic & Hamilton, 2013) and voices (e.g. Globerson, Amir, Kishon-Rabin, & Golan, 2015; Hobson, Ouston, & Lee, 1988; Stewart, McAdam, Ota, Peppe, & Cleland, 2013). Pellicano, Jeffery, Burr and Rhodes (2007) have suggested that one underlying cause of face perception difficulties may be an impaired adaptation mechanism.

Adaptation is thought to be a universal mechanism across all sensory systems (Møller, 2002) and refers to a change in neural firing and therefore behaviour as a result of continued stimulation or exposure to the same sensory stimulus (Quiroga & Panzeri, 2013). The generally accepted view is that adaptation is essential for efficient coding and processing of stimuli by allowing novel features to appear more salient (Webster, 2011). Research using adaptation has revealed neural populations tuned to respond to specific stimulus attributes by isolating and subsequently distorting the perception of these attributes (Grill-Spector et al., 1999; Winston, Henson, Fine-Goulden, & Dolan, 2004). Adaptation has therefore been a useful tool to investigate different levels of processing including “low-level” stimulus properties such as motion, shape or colour (e.g. Wright, 1934) but also “higher-level” properties such as facial or vocal features (e.g. emotional expressions).

Webster and MacLin (1999) were the first to show that extended exposure to faces can generate aftereffects, i.e. a perceptual bias in which the test face was interpreted as being distorted in the opposite direction to that of the adaptor face. Adaptation to consistently distorted faces (e.g. expanded features) caused subsequently viewed unmanipulated faces to

appear distorted in the opposite direction of the adapting stimulus (e.g. compressed features). This effect transferred to faces of different identities. Later studies demonstrated that these contrastive aftereffects are robust to changes in the retinal position (Leopold, O'Toole, Vetter, & Blanz, 2001), size (Zhao & Chubb, 2001) and angular orientation (Watson & Clifford, 2003) of the adapting stimulus, suggesting that adaptation to those low-level features cannot fully explain face aftereffects. Thus, adaptive coding is a strategy employed by the neural system throughout the perceptual processing hierarchy and is not only employed at early stages of visual information processing. Moreover, even high-level, socially important properties such as facial identity (Leopold et al., 2001), attractiveness (Rhodes, Jeffery, Watson, Clifford, & Nakayama, 2003), eye-gaze (Jenkins et al., 2006) and expression (Fox & Barton, 2007; Webster et al., 2004) can be isolated using adaptation. Researchers have interpreted the resulting aftereffects as reflecting a neural strategy for optimising perception, including a recalibration of neural processes to continuously updated stimulation (Webster & MacLin, 1999). While research has used these contrastive aftereffects as a tool to evaluate face perception models (e.g. Bestelmeyer et al., 2008; Bestelmeyer, Jones, DeBruine, Little, & Welling, 2010; Jones, Feinberg, Bestelmeyer, DeBruine, & Little, 2010; Leopold, O'Toole, Vetter, & Blanz, 2001; Leopold, Rhodes, Muller, & Jeffery, 2005; Little, DeBruine, & Jones, 2005; Webster, Kaping, Mizokami, & Duhamel, 2004; Welling et al., 2009), it has recently been shown that corresponding auditory aftereffects can also be elicited for voices (e.g. identity: Zaska, Schweinberger, & Kawahara, 2010; gender: Schweinberger et al., 2008; age: Zaska & Schweinberger, 2011; and affect: Bestelmeyer, Rouger, DeBruine, & Belin, 2010; Skuk & Schweinberger, 2013).

In an influential paper, Pellicano et al. (2007) reported that children with ASD show significantly reduced facial identity aftereffects compared to IQ-matched, typically developing children. While face identification was actually unimpaired in both groups the authors suggested that an impaired adaptive face-coding mechanism in ASD may be responsible for the commonly reported face perception problems as well as some of the core social deficits

in ASD. A follow-up study ascertained that these reduced aftereffects in ASD are not caused by a lack of attention to the adapting stimulus (Ewing, Leach, Pellicano, Jeffery, & Rhodes, 2013) supporting that children with ASD have difficulties with adaptive norm-based coding of face identity. Siblings and parents of children with ASD also show this reduction in aftereffect size to identity adaptation suggesting that this behavioural difference may have the potential to serve as an endophenotype for ASD (Fiorentini, Gray, Rhodes, Jeffery, & Pellicano, 2012). Recent evidence of reduced adaptation despite intact recognition of biological motion (van Boxtel, Dapretto, & Lu, 2016), another social stimulus, further strengthens the possibility that changes in the adaptation mechanism contribute to the social phenotype in ASD.

The literature is not entirely unequivocal regarding reduced aftereffects to social stimuli in ASD. For example, recent studies have shown that reduced aftereffects can also be observed to non-social stimuli, e.g. reduced aftereffects to numerosity (Turi et al., 2015) and audio-visual asynchrony (Turi, Karaminis, Pellicano & Burr, 2016), although not consistently to all non-social stimuli, e.g. intact adaptation to perceptual causality (Karaminis et al., 2015). Similarly, it is interesting that many of the reduced face aftereffects found in children with ASD cannot be replicated in adults with ASD (e.g. Cook, Brewer, Shah, & Bird, 2014; Walsh, Maurer, et al., 2015; Walsh, Vida, Morrissey, & Rutherford, 2015). In isolation, this last pattern of results would suggest that reduced contrastive aftereffects for faces in children with ASD reflect delayed or atypical development of adaptive mechanisms and as such may not be a core feature of the disorder. However, this notion is in contrast to Fiorentini et al.'s (2012) finding of reduced visual aftereffects in relatives (including adult relatives) of ASD children. If future work confirms that reduced face aftereffects are merely due to developmental delays, they are unlikely to be specific to autism and may not be useful as an endophenotype of ASD.

In light of these findings, several questions remain unanswered regarding the specificity of reduced or sometimes even absent contrastive aftereffects in ASD. First, it is unclear from the current literature whether this alteration in adaptation strength is specific to the stimulus type (e.g. social or non-social), developmental stage or, indeed, modality. Only

one paper has explored adaptation in ASD in the auditory domain. Lawson, Aylward, White, and Rees (2015) showed that continuous stimulation with sound of a certain volume caused a reduced magnitude and rate of simple loudness adaptation in ASD but not in typically developing children. While this finding theoretically provides support for a generalized reduction in adaptation in ASD, it may instead be in line with increased sound sensitivity in ASD and does not rule out the possibility that adaptive mechanisms to more complex sound categories (e.g. social stimuli) are intact. Thus whether children with ASD also show reduced *contrastive* aftereffects in response to adaptation using more complex auditory stimuli has not been determined. The advantage of using auditory over visual stimuli to assess the specificity of the reduced aftereffect to certain stimulus types (social vs non-social) is that even distractible children cannot ignore auditory stimuli presented via headphones as easily as visually presented stimuli on a small computer display.

The aim of the current study was to address some of these remaining questions. First, we examined whether reduced contrastive aftereffects in ASD also extend to high-level stimuli in the auditory domain. Second, we assessed whether aftereffects are smaller or even absent in a task with emotionally salient stimuli compared to a control task in ASD. More specifically, we were interested in group differences in the perceptual plasticity of two vocal features with differing emotional content, which could form part of the phenotype of ASD. Using auditory voice morphing we created blends between two instances of two stimulus categories; emotional vocalisations (anger and fear) and phonemes ('m' and 'o'). Given the previous literature, we predicted both a main effect of participant group and a significant interaction between stimulus type and group when measuring the magnitudes of contrastive aftereffects. In other words, we expected a reduction in aftereffects across all stimuli in the ASD group as well as a greater aftereffect reduction in ASD to the more socially salient emotion stimuli, compared to the phoneme stimuli. We expected the typically developing children to show robust aftereffects across condition and be unaffected by stimulus type.

Methods

Participants

Based on the between group difference reported by Pellicano et al. (2007), we performed a power analysis in G*power 3 (Faul, Erdfelder, Buchner, & Lang, 2009). This analysis determined that at an alpha level of $\alpha = .05$ (2-tailed), an effect size of $d = .81$ and a power of $\beta = .8$ a sample size of 25 participants is required in each group. We therefore recruited a total of 30 Caucasian children with ASD as outpatients of the Department of Child and Adolescent Psychiatry and Psychotherapy, Medical Faculty, University of Freiburg, along with 26 age- and culture-matched, typically developing controls. Due to incomplete data sets or misunderstandings of the task, we had to exclude 3 ASD and 2 typically developing children. Our analyses are therefore based on 27 (2 female) children with ASD and 24 (8 female) typically developing children. Table 1 summarises age and IQ as measured by the Cattell Culture Fair Intelligence Test (CFT-2R; Weiss, 2006) for the included participants of both groups except for one of the children with ASD where IQ was assessed using the Snijders-Oomen nonverbal IQ test (Tellegen, Winkel, Wijnberg-Williams & Laros, 1998). His score was 78 and is not included in Table 1. Children with ASD had a diagnosis according to ICD-10 and DSM IV criteria. Diagnoses of ASD are based on the international standard, including the German versions of the Autism Diagnostic Observation Schedule (ADOS; Rühl, Bölte, Feineis-Matthews, & Poustka, 2004), the Social Responsiveness Scale (SRS; Bölte & Poustka, 2007) and of the Autism Diagnostic Interview (ADI-R; Bölte, Rühl, Schmötzer, & Poustka, 2006) which is a semi-structured interview of the caregiver and their observations of their child. Table 1 provides means, standard deviations and ranges of the social, communication and total ADOS score, the SRS raw and T-Norms as well as the three categories indicating autistic behaviour of the ADI-R (“social interaction”, “social communication” and “restricted, repetitive, stereotyped behaviours”) for the children with ASD. In our sample, all participants but one were above cut-off in the ADI-R category “social interaction”, only one was under the cut-off for “communication”, four participants did not reach

the cut-off for “restricted, repetitive, stereotyped behaviors”. Overall, all patients exceeded cut-off at least in two domains of the ADI-R (which is a criterion for a diagnosis of ASD) and also scored in the third “autistic” domain (although often under cut-off). Moreover, in the ADOS all patients but two (who scored slightly under cut-off) received an ADOS-diagnosis of autism spectrum or autism. These two patients were above cut-off at least in two domains of the ADI-R. Only one participant (not previously mentioned) scored under cut-off for deficits of social responsiveness.

Our sample is high-functioning with good cognitive and communication abilities and the severity of symptoms can be considered at level 2 (middle) according to DSM-V. The symptom severity of our sample is comparable to that of Pellicano et al. (2007). This study was approved by the ethics committee of Bangor University for piloting with healthy adults and of Freiburg University for the testing done with the children. Parents of all children gave written informed consent at the beginning of the test session.

TABLE 1 HERE

Stimuli

Recordings of the emotional sounds were taken from the Montreal Affective Voices database (Belin, Fillion-Bilodeau & Gosselin, 2008) in which actors were instructed to produce emotional interjections using the vowel /a/. The voices from four identities (two female) expressed anger and fear. Recordings of the phonemes (/m/ and /o/) were made in a double-walled sound attenuated booth at Bangor University using again four different German native speakers (two female). The phoneme recordings were edited using CoolEditPro 2.0 (Syntrillium, Inc.) to single utterances of ‘m’ and ‘o’ sounds (i.e. editing was comparable to that of the emotional expressions in Belin et al. (2008)). All sounds were normalised in energy (root mean square).

We created morph continua separately for each identity in the emotion and in the phoneme condition using Tandem STRAIGHT (Banno et al., 2007; Kawahara et al., 2008; see also http://www.wakayama-u.ac.jp/~kawahara/STRAIGHTadv/index_e.html for more information). Each continuum consisted of five steps corresponding to 5/95%, 35/65%, 50/50%, 65/35% and 95/5% fear/anger or 'm'/'o'. The original angry (0/100%) and fearful (100/0%) voices were used as adaptors in the emotion task and the original 'm' (0/100%) and 'o' (100/0%) sounds were used as adaptors in the phoneme task. Tandem STRAIGHT performs an instantaneous pitch-adaptive spectral smoothing of each stimulus for separation of contributions to the voice signal arising from the glottal source (including f₀) versus supralaryngeal filtering (distribution of spectral peaks, including the first formant frequency, F₁). Voice stimuli were decomposed by Tandem STRAIGHT into five parameters: f₀ (the perceived pitch of the voice), frequency, duration, spectrotemporal density and aperiodicity. Each parameter can be manipulated independently. For each voice we manually identified one time landmark with three frequency landmarks (corresponding to the first three formants) at the onset of phonation and the same number of landmarks at the offset of phonation. Morphed stimuli were then generated by re-synthesis based on the interpolation (linear for time; logarithmic for F₀, frequency, and amplitude) of these time-frequency landmark templates (see also Schweinberger et al. (2014) for a recent discussion of the voice morphing technique). Stimuli were normalised in energy (root mean square) before and after morphing. Acoustic analyses to illustrate the validity of the morphing technique with affective bursts, along with sample stimuli, are published elsewhere (Bestelmeyer et al., 2010). The duration of the vocalisations within each continuum was kept constant and ranged between 0.6 and 0.9s across continua. We used E-prime2 (Psychology Software Tools, Inc.) for stimulus presentation. Sounds were presented via headphones at a comfortable sound level for the child.

To insure that stimulus categorisations for both emotion and phoneme tasks would yield similar sigmoid shaped curves with comparable points of subjective equality and slope

we first tested both tasks on 57 young, healthy undergraduate students (35 female; mean age = 20.86; S.D. = 5.19). The slopes of the curves for each task were not significantly different ($t(56) = -1.00, p = .322$). Both tasks resulted in a significant main effect of adaptation (emotion task: $F(2, 112) = 79.261, p < .001$; speech task: $F(2, 112) = 147.64, p < .001$). The emotion stimuli are the same as the ones used in several other studies for which categorisation consistently resulted in sigmoid shaped curves (Bestelmeyer, Maurage, Rouger, Latinus, & Belin, 2014; Bestelmeyer et al., 2010; Pye & Bestelmeyer, 2015) and robust contrastive aftereffects in adults.

Procedure

Participants were asked to complete a total of five blocks for each task. All tasks required a two-alternative forced choice judgement. The emotion task required a judgment as to whether the ambiguous vocal morph sounded more angry or fearful and the phoneme task required a judgment as to whether the vocal morph sounded more like the letter 'm' or 'o'. Each task consisted of one pre-adaptation or baseline block, which was always administered before the four adaptation blocks of each task. The trials of the baseline block consisted of individual morphs that had to be categorised. Each baseline block was comprised of three repetitions of each of the five morph steps per identity (2 male, 2 female) leading to a total of 60 trials. Morphs were presented in randomised order but blocked for gender.

The trial structure of the adaptation tasks consisted of one adapting voice played four times in succession and followed by an ambiguous morph (test stimulus) after a silent gap of one second. The inter-trial interval was response-terminated at a maximum of ten seconds long before the next trial would commence. During this interval the participants categorised the test stimulus using one of two keys ("Does the person sound angry or fearful"/"Does the person say 'M' or 'O'"). Each of the four adaptation blocks (2 emotion x 2 gender or 2 phoneme x 2 gender) and each of the five test morphs per identity were repeated six times leading to a

total of 30 trials per block. Trials were presented in randomised order but blocked for task type and adaptation condition (i.e. both blocks consisting of female and male fear adaptation were presented in succession rather than mixing the emotion adaptation conditions).

To avoid adaptation to task irrelevant factors such as identity-specific or low-level acoustic properties, participants were always tested on a different identity than the one they were adapted to. Half of the participants were adapted to angry and then fearful voices while the remaining participants were presented with the reverse order of blocks. The same protocol was used for the phoneme adaptation task. Task order was also counterbalanced across participants. The raw data for this experiment is published with Mendeley (doi:10.17632/dzcszhwsfp.1).

Results

For each participant, responses were averaged as a function of the five morph steps, adaptation condition and task. We then fitted the data with a logistic function with four free parameters: the minimum and maximum x- and y-values at the centre of symmetry of the curve (the point of subjective equality (PSE)) and the slope of the curve at the PSE. Excellent fits were obtained for both tasks across all conditions for each participant group (emotion task for the ASD group: mean $R^2 = .926$, SD = .186 and the TD group: mean $R^2 = .956$, SD = .123; phoneme task for the ASD group: mean $R^2 = .987$, SD = .027 and the TD group mean $R^2 = .992$, SD = .013).

Figure 1 illustrates the psychometric curves fitted to the mean data for the participant groups and tasks. Figure 1 A and D illustrate the mean responses of the emotion and phoneme baseline conditions (no adaptation), respectively. Figure 1 B and E illustrate the mean responses of the adaptation conditions to emotions and phonemes, respectively. The group PSE is illustrated with a star on all average curves in corresponding colour. The slopes in the baseline tasks are not significantly different, suggesting similar difficulty levels across tasks.

Aftereffect sizes do not differ across participant groups. The bar graphs in Figure 1 C and F show the aftereffect sizes (condition A – condition B at the PSE) for the emotion and phoneme tasks, respectively, as well as individual participant's data (black circles).

FIGURE 1 HERE

To ensure that the tasks did not differ in difficulty we calculated the slope for the baseline curves. Using a mixed design ANOVA (2 groups x 2 tasks) we found no significant main effects of task ($F(1, 49) = .503$; $p = .482$; $\eta^2 = .010$) or group ($F(1, 49) = .524$; $p = .472$; $\eta^2 = .011$) and no significant interaction between group and task ($F(1, 49) = .217$; $p = .643$; $\eta^2 = .004$) on the slope of the curves. Interestingly, from Figure 1A, it is obvious that the extreme emotion categories were less accurately categorised in the children with ASD than the typically developing children (one-way ANOVA for the two extreme morph steps; anger: $F(1, 49) = 5.124$; $p = .028$ and fear: $F(1, 49) = 5.894$; $p = .019$). This pattern was not evident in the control task (Figure 1B; both $F(1, 49) < 1.079$; $p > .304$).

To assess any group differences in aftereffect size in each task and adaptation condition (including the baseline condition) we carried out a mixed design ANOVA (2 participant groups x 2 tasks x 3 adaptation conditions) on the PSE abscissa values, which revealed a significant main effect of adaptation condition ($F(2, 98) = 151.821$; $p < .001$; $\eta^2 = .756$). Pairwise comparisons revealed significant differences between all adaptation conditions, including the baseline task (all at $p < .0001$). Contrary to our hypotheses, there was no significant main effect of group ($F(1, 49) = .699$; $p = .407$; $\eta^2 = .014$) and no significant interaction between group and task ($F(1, 49) = .326$; $p = .571$; $\eta^2 = .007$). We did find a significant interaction between task and adaptation condition, which was driven by a significantly larger aftereffect to the phoneme sounds than to the vocal expressions. There were no other significant main effects or interactions.

We also examined the possibility that children with ASD adapt more slowly by testing for aftereffect differences between the two groups in only the first half of trials. This analysis resulted in the same pattern as the analysis on the full set of trials (i.e. no main effect of group or significant interaction between group and task). We also explored the possibility that variables such as age, IQ and AQ might co-vary with our effects of interest but none of these variables did (all $F < 1.6$; $p > .22$). Similarly, none of these variables significantly correlated with the aftereffect size (PSE of fear – PSE of anger adaptation; PSE of O adaptation – PSE of M adaptation; all $r < -.19$; $p > .19$).

Discussion

The current study tested two key questions: first, will differences in contrastive aftereffects previously shown using visual stimuli in children with ASD be evident within the auditory domain; and second, will such differences be seen only in response to the more socially salient stimuli? Recent influential work has suggested that changes in adaptive mechanisms in ASD may be at the root of some of the core aspects of the social phenotype in ASD, including deficits in social cognition and social interaction. However, using auditory stimuli, our within-subjects study revealed no differences in aftereffect sizes for the two participant groups in either the more socially salient or the well-matched control stimuli. In other words, children with ASD did not show any reductions in aftereffect sizes compared to typically developing children regardless of auditory stimulus type. We also found no suggestions of a slowed adaptation rate in ASD compared to typically developing children. It is important to note here that this is not simply a null finding; instead, we report robust adaptation effects that are independently significant in each group and virtually identical in magnitude across groups.

Several research groups have reported that the mere categorisation or recognition of social stimuli is impaired in ASD (e.g. emotional faces: Harms et al., 2010; Teunisse & de

Gelder, 2001; emotional faces and speech prosody: Globerson, Amir, Kishon-Rabin, & Golan, 2015; Philip et al., 2010; Taylor, Maybery, Grayndler, & Whitehouse, 2015). Interestingly, we observed a similar pattern in the ASD group to the socially salient vocal expressions of anger and fear endpoints in the baseline categorisation task confirming limitations in recognizing emotional expressions also in the auditory domain. In a seminal paper, Pellicano et al. (2007) suggested that this poorer performance in categorising social stimuli may be caused by, or may at least be related to, impairments in the adaptive mechanism in ASD. Our data do not support this claim. While we replicate previous findings of poorer categorisation performance of the emotional expressions in the ASD group, we find no group differences in the magnitude of the aftereffects to emotional expressions. This pattern of results suggests that impaired recognition of emotions but intact perceptual plasticity in ASD are unlikely to be causally linked.

There were several differences in terms of task design and stimulus type between Pellicano et al.'s study and the present one which preclude a definite answer as to whether adaptive coding plays a role in the cognitive phenotype of ASD. While both studies employed social stimuli (facial identity, vocal emotion) our paradigm was not designed around an arguably more complex game. Rather, our study was based on a traditional cognitive adaptation paradigm with four repetitions of the adaptor followed by a test stimulus that is known to yield robust adaptation effects. While our paradigm may have been less engaging because it was not embedded in a story, it is possible that at least some children, particularly those with ASD, might find our paradigm easier to deal with because no social understanding or involvement was required to complete the task. More specifically, it is possible that the task used by Pellicano et al. was actually less suitable for children with ASD due to the social framing of the task which might put children with ASD at a disadvantage. Whether the discrepancy in results of the two studies is due to differences in modality or task therefore remains an open question because later face adaptation studies in ASD also typically use tasks designed around a game or story.

Lawson et al. (2015) investigated two types of low-level auditory adaptation in adults with ASD. First, participants had to repeatedly judge the loudness of a quiet, continuously and monaurally presented sound. This type of adaptation (simple loudness) was impaired in adults with ASD compared to healthy controls and may give rise to increased sound sensitivity in ASD. Second, participants had to judge the volume of a much louder, continuously presented sound while a distracter sound was intermittently played to the other ear. This type of adaptation (induced loudness) was intact in the autism group. The authors point out that different neural substrates likely underpin each adaptation type but it is unclear why the participants with ASD were impaired for simple but not induced loudness adaptation. The results do suggest, however, that the basic adaptation mechanism is not impaired in ASD.

Studies following Pellicano's seminal paper have not been entirely unequivocal regarding reduced aftereffects to social stimuli in ASD. Most studies have shown diminished aftereffect sizes (Ewing et al., 2013; Pellicano et al., 2007; Pellicano, Rhodes, & Calder, 2013; Turi et al., 2015; 2016), one recent study has shown a priming effect in ASD but not the control group (Rutherford, Troubridge, & Walsh, 2012) and only a few studies, on adult participants with ASD, have found no group differences in the face adaptation effect size (Cook et al., 2014; Walsh, Maurer, et al., 2015; Walsh, Vida, Morrisey, & Rutherford, 2015). Participant numbers (or statistical power) are unlikely to explain the differences between previous results and results of the current study as participant numbers in the current study are larger than in many of the aforementioned reports (all but Ewing et al., 2013 and Walsh, Maurer et al., 2015). There are several possible explanations for these discrepancies. For example, while adaptation is a ubiquitous and fundamental mechanism across sensory modalities, it is possible that differences or developmental delays in *visual* adaptation are modality specific and do not affect adaptation to high-level auditory stimuli. Equally, it is plausible that auditory stimuli are less prone to attentional lapses than visual stimuli and that, despite attempts to control attention in visual paradigms (Ewing et al., 2013), reduced attention to the adapter stimuli in ASD does play a role, particularly in the child population. Importantly, if atypical

adaptation to faces is only seen in *children* with ASD, then the effect may be due to delayed or atypical development and therefore would not be a stable aspect of the condition. Additionally, as our study demonstrates typical adaptation aftereffects to a “social” (albeit auditory) stimulus in childhood, atypical adaptation in ASD also cannot be assumed in response to all social stimuli. Thus the suitability of impaired adaptation as an endophenotype is unlikely. At this point, smaller aftereffects demonstrated in some studies but not others may be due to a myriad of factors but it cannot be seen as a robust or universal impairment. We can therefore exclude impaired aftereffects as reliably descriptive of the cognitive phenotype of ASD.

Our results highlight the need for a more systematic investigation of adaptation in autism to account for the inconsistent findings in the literature. To understand the nature of the putative impairment of the adaptation mechanism in ASD, a thorough investigation of contrastive aftereffects using different stimulus types (social, non-social) at different levels of the processing hierarchy (low-level vs high level), for various modalities and across different stages of development is required.

Conclusion

We show for the first time that auditory contrastive aftereffects in children with ASD are intact for both socially salient stimuli and control stimuli thereby challenging the widely-held belief that adaptation mechanisms are specifically impaired in autism spectrum disorders. While children with ASD were significantly worse at categorising the endpoints of the vocal emotion continua during the baseline task, our results demonstrate that these differences are not directly related to the functioning of the adaptation mechanism. When using a traditional adaptation paradigm with a categorisation task on the high-level dimensions of sound, the adaptation mechanism in ASD is intact.

References

- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders*. (5th ed ed.). Arlington, VA: American Psychiatric Publishing.
- Banno, H., Hata, H., Morise, M., Takahashi, T., Irino, T., & Kawahara, H. (2007). Implementation of realtime STRAIGHT manipulation system: Report on its first implementation. *Acoustical Science and Technology*, *28*, 140-146.
- 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.
- Bestelmeyer, P. E. G., Jones, B. C., DeBruine, L. M., Little, A. C., Perrett, D. I., Schneider, A., . . . Conway, C. A. (2008). Sex-contingent face aftereffects depend on perceptual category rather than structural encoding. *Cognition*, *107*(1), 353-365
- Bestelmeyer, P. E. G., Jones, B. C., DeBruine, L. M., Little, A. C., & Welling, L. L. M. (2010). Face aftereffects demonstrate interdependent processing of expression and sex and of expression and race. *Visual Cognition*, *18*, 255-274.
- Bestelmeyer, P. E. G., Maurage, P., Rouger, J., Latinus, M., & Belin, P. (2014). Adaptation to Vocal Expressions Reveals Multistep Perception of Auditory Emotion. *Journal of Neuroscience*, *34*(24), 8098-8105. doi:10.1523/jneurosci.4820-13.2014
- Bestelmeyer, P. E. G., Rouger, J., DeBruine, L. M., & Belin, P. (2010). Auditory adaptation in vocal affect perception. *Cognition*, *117*, 217-223.
- Bölte, S., & Poustka, F. (2007). *Skala zur Erfassung sozialer Reaktivität—Dimensionale Autismus-Diagnostik. German version of the social responsiveness scale (SRS) by John N. Constantino und Christian P. Gruber*. Bern: Huber.
- Bölte, S., Rühl, D., Schmötzer, G., & Poustka, F. (2006). *Diagnostisches Interview für Autismus - Revidiert (ADI-R)*. Bern: Huber.
- Cook, R., Brewer, R., Shah, P., & Bird, G. (2014). Intact Facial Adaptation in Autistic Adults. *Autism Research*, *7*(4), 481-490. doi:10.1002/aur.1381

- Ewing, L., Leach, K., Pellicano, E., Jeffery, L., & Rhodes, G. (2013). Reduced Face Aftereffects in Autism Are Not Due to Poor Attention. *Plos One*, 8(11). doi:10.1371/journal.pone.0081353
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149–1160.
- Fiorentini, C., Gray, L., Rhodes, G., Jeffery, L., & Pellicano, E. (2012). Reduced face identity aftereffects in relatives of children with autism. *Neuropsychologia*, 50(12), 2926-2932. doi:10.1016/j.neuropsychologia.2012.08.019
- Globerson, E., Amir, N., Kishon-Rabin, L., & Golan, O. (2015). Prosody Recognition in Adults With High-Functioning Autism Spectrum Disorders: From Psychoacoustics to Cognition. *Autism Research*, 8(2), 153-163. doi:10.1002/aur.1432
- Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron*, 24(1), 187-203.
- 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. doi:10.1007/s11065-010-9138-6
- Hobson, R. P., Ouston, J., & Lee, A. (1988). Emotion recognition in autism - coordinating faces and voices. *Psychological Medicine*, 18(4), 911-923.
- Jones, B. C., Feinberg, D. R., Bestelmeyer, P. E. G., DeBruine, L. M., & Little, A. C. (2010). Adaptation to different mouth shapes influences visual perception of ambiguous lip speech. *Psychonomic Bulletin & Review*, 17(4), 522-528.
- Karaminis, T., Turi, M., Neil, L., Badcock, N. A., Burr, D., & Pellicano, E. (2015). Atypicalities in Perceptual Adaptation in Autism Do Not Extend to Perceptual Causality. *Plos One*, 10(3). doi:10.1371/journal.pone.0120439
- Kawahara, H., Morise, M., Takahashi, T., Nisimura, R., Irino, T., & Banno, H. (2008). *Tandem-STRAIGHT: A temporally stable power spectral representation for periodic signals and*

applications to interference-free spectrum, F0, and aperiodicity estimation. Paper presented at the ICASSP Las Vegas.

- Lawson, R. P., Aylward, J., White, S., & Rees, G. (2015). A striking reduction of simple loudness adaptation in autism. *Scientific Reports*, 5. doi:10.1038/srep16157
- Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level after effects. *Nature Neuroscience*, 4(1), 89-94.
- Leopold, D. A., Rhodes, G., Muller, K. M., & Jeffery, L. (2005). The dynamics of visual adaptation to faces. *Proceedings of the Royal Society B-Biological Sciences*, 272(1566), 897-904.
- Little, A. C., DeBruine, L. M., & Jones, B. C. (2005). Sex-contingent face after-effects suggest distinct neural populations code male and female faces. *Proceedings of the Royal Society B-Biological Sciences*, 272(1578), 2283-2287.
- Marco, E. J., Hinkley, L. B. N., Hill, S. S., & Nagarajan, S. S. (2011). Sensory Processing in Autism: A Review of Neurophysiologic Findings. *Pediatric Research*, 69, 48R-54R.
- Møller, A.R. (2002). *Sensory Systems: Anatomy and Physiology*. San Diego: Academic Press.
- Pellicano, E., Jeffery, L., Burr, D., & Rhodes, G. (2007). Abnormal adaptive face-coding mechanisms in children with autism spectrum disorder. *Current Biology*, 17(17), 1508-1512. doi:10.1016/j.cub.2007.07.065
- Pellicano, E., Rhodes, G., & Calder, A. J. (2013). Reduced gaze aftereffects are related to difficulties categorising gaze direction in children with autism. *Neuropsychologia*, 51(8), 1504-1509.
- Philip, R. C. M., Whalley, H. C., Stanfield, A. C., Sprengelmeyer, R., Santos, I. M., Young, A. W., Atkinson, A. P., Calder, A. J., Johnstone, E. C., Lawrie, S. M., & Hall, J. (2010). *Psychological Medicine*, 40, 1919-1929.
- Psychology Software Tools, Inc. [E-Prime 2.0]. (2012). Retrieved from <http://www.pstnet.com>.
- Pye, A., & Bestelmeyer, P. E. G. (2015). Evidence for a supra-modal representation of emotion from cross-modal adaptation. *Cognition*, 134, 245-251. doi:10.1016/j.cognition.2014.11.001

- Quiroga, R. Q. & Panzeri, S. (2013). Principles of Neural Coding. CRC Press.
- Rühl, D., Bölte, S., Feineis-Matthews, S., & Poustka, F. (2004). *Diagnostische Beobachtungsskala für Autistische Störungen (ADOS)*. Bern: Huber.
- Rutherford, M. D., Troubridge, E. K., & Walsh, J. (2012). Visual Afterimages of Emotional Faces in High Functioning Autism. *Journal of Autism and Developmental Disorders*, 42(2), 221-229. doi:10.1007/s10803-011-1233-x
- Schweinberger, S. R., Casper, C., Hauthal, N., Kaufmann, J. M., Kawahara, H., Kloth, N., . . . Zäske, R. (2008). Auditory adaptation in voice perception. *Current Biology*, 18(9), 684-688.
- Schweinberger, S.R., Kawahara, H., Simpson, A.P., Skuk, V.G., Zäske, R. (2014). Speaker perception. *WIREs Cogn Sci*2014, 5:15–25. doi: 10.1002/wcs.1261
- Skuk, V. G., & Schweinberger, S. R. (2013). Adaptation Aftereffects in Vocal Emotion Perception Elicited by Expressive Faces and Voices. *Plos One*, 8(11). doi:10.1371/journal.pone.0081691
- Stewart, M. E., McAdam, C., Ota, M., Peppe, S., & Cleland, J. (2013). Emotional recognition in autism spectrum conditions from voices and faces. *Autism*, 17(1), 6-14. doi:10.1177/1362361311424572
- Taylor, L. J., Maybery, M. T., Grayndler, L., & Whitehouse, J.O. (2015). Evidence for shared deficits in identifying emotions from faces and from voices in autism spectrum disorders and specific language impairment. *International Journal of Language and Communication Disorders*, 50(4), 452-466.
- Tellegen, P.J., Winkel, M., Wijnberg-Williams, B. & Laros, J.A. (1998). SON-R 2,5-7: Manual and research report. Amsterdam: Hogrefe uitgevers.
- Teunisse, J.-P. & de Gelder, B. (2001). Impaired Categorical Perception of Facial Expressions in High-Functioning Adolescents with Autism. *Child Neuropsychology*, 7(1), 1-14.
- Turi, M., Burr, D. C., Igliozzi, R., Aagten-Murphy, D., Muratori, F., & Pellicano, E. (2015). Children with autism spectrum disorder show reduced adaptation to number.

- Proceedings of the National Academy of Sciences of the United States of America*, 112(25), 7868-7872. doi:10.1073/pnas.1504099112
- Turi, M., Karaminis, T., Pellicano, E., & Burr, D. (2015). No rapid audiovisual recalibration in adults on the autism spectrum. *Scientific Reports*, 6, 21756. doi: 10.1038/srep21756.
- Uljarevic M, Hamilton A. Recognition of emotions in autism: a formal meta-analysis. *Journal of Autism and Developmental Disorders*, 43(7),1517-1526. doi: 10.1007/s10803-012-1695-5.
- van Boxtel, J. J., Dapretto, M. & Lu, H. (2016). Intact recognition, but attenuated adaptation, for biological motion in youth with autism spectrum disorder. *Autism Research*, 9(10), 1103-1113. doi: 10.1002/aur.1595
- Walsh, J. A., Maurer, D., Vida, M. D., Rhodes, G., Jeffery, L., & Rutherford, M. D. (2015). Norm-based coding of facial identity in adults with autism spectrum disorder. *Vision Research*, 108, 33-40. doi:10.1016/j.visres.2014.11.018
- Walsh, J. A., Vida, M. D., Morrisey, M. N., & Rutherford, M. D. (2015). Adults with autism spectrum disorder show evidence of figural aftereffects with male and female faces. *Vision Research*, 115, 104-112. doi:10.1016/j.visres.2015.08.010
- Webster, M. A. (2011). Adaptation and visual coding. *Journal of Vision*, 11, 10.1167. doi: 10.1167/11.5.3
- Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature*, 428(6982), 557-561.
- Webster, M. A., & MacLin, O. H. (1999). Figural aftereffects in the perception of faces. *Psychonomic Bulletin & Review*, 6(4), 647-653.
- Weigelt, S., Koldewyn, K., & Kanwisher, N. (2012). Face identity recognition in autism spectrum disorders: A review of behavioral studies. *Neuroscience and Biobehavioral Reviews*, 36(3), 1060-1084. doi:10.1016/j.neubiorev.2011.12.008
- Weiss, R. H. (2006). *Grundintelligenztest Skala 2 (CFT 20-R)*. Goettingen: Hoegrefe.

- Welling, L. L. M., Jones, B. C., Bestelmeyer, P. E. G., DeBruine, L. M., Little, A. C., & Conway, C. A. (2009). View-contingent aftereffects suggest joint coding of face shape and view. *Perception, 38*(1), 133-141.
- Winston, J. S., Henson, R. N. A., Fine-Goulden, M. R., & Dolan, R. J. (2004). fMRI-adaptation reveals dissociable neural representations of identity and expression in face perception. *Journal of Neurophysiology, 92*(3), 1830-1839. doi:10.1152/jn.00155.2004
- Wright, W. D. (1934). The measurement and analysis of colour adaptation phenomena. *Proceedings of the Royal Society of London. Series B, 115*, 49-87.
- Zäske, R., & Schweinberger, S. R. (2011). You are only as old as you sound: Auditory aftereffects in vocal age perception. *Hearing Research, 282*(1-2), 283-288. doi:10.1016/j.heares.2011.06.008
- Zäske, R., Schweinberger, S. R., & Kawahara, H. (2010). Voice aftereffects of adaptation to speaker identity. *Hearing Research, 268*(1-2), 38-45. doi:10.1016/j.heares.2010.04.011

Table 1. Participant demographics and test scores.

| Measure | Children with autism | | | | Typically developing children | | | | t | p-value |
|------------------------------|----------------------|-------|-------|--------------|-------------------------------|--------|-------|--------------|-------|---------|
| | N | Mean | SD | Range | N | Mean | SD | Range | | |
| Age | 27 | 11.87 | 2.38 | 8.10 - 17.69 | 24 | 12.20 | 2.14 | 6.46 - 15.22 | -0.52 | 0.607 |
| IQ (CFT 20-R) | 25 | 96.00 | 14.02 | 71 - 122 | 24 | 109.08 | 14.89 | 84 - 141 | -3.26 | 0.002 |
| SRS Raw Score | 27 | 99.48 | 33.82 | 30 - 170 | | | | | | |
| SRS T Norms | 27 | 80.89 | 11.79 | 55 - 100 | | | | | | |
| ADOS (social) | 27 | 7.85 | 1.83 | 4 - 11 | | | | | | |
| ADOS (communication) | 27 | 3.89 | 2.04 | 0 - 8 | | | | | | |
| ADOS (total) | 27 | 11.74 | 3.05 | 6 - 17 | | | | | | |
| ADI-R (social) | 26 | 18.58 | 5.14 | 10 - 26 | | | | | | |
| ADI-R (communication) | 26 | 12.77 | 5.18 | 1 - 21 | | | | | | |
| ADI-R (repetitive behaviour) | 26 | 5.12 | 3.34 | 0 - 13 | | | | | | |

Figure 1. Psychometric functions were fitted for the mean responses to each morph step and separately for each task and participant group. Solid lines always indicate the performance of the children with ASD; dashed lines always indicate the performance of the typically developing children (TDC). The point of subjective equality (PSE) is plotted as a star on each curve (stars for the PSEs of the ASD children are filled; stars for the PSEs of the TDC children are white). Baseline tasks for the emotion and phoneme task are plotted in A and D, respectively. Adaptation tasks for the emotion and phoneme task are plotted in B and E. Error bars represent standard error of the mean. C and F show the aftereffect sizes for each participant group based on the PSE in each condition as well as individual participant's data (black circles).

