Psychological Research
https://doi.org/10.1007/s00426-018-1133-1

ORIGINAL ARTICLE

The representation and plasticity of body emotion expression

Rebecca Watson[1] · Beatrice de Gelder[1]

Received: 22 October 2015 / Accepted: 10 December 2018

Abstract

Emotions are expressed by the face, the voice and the whole body. Research on the face and the voice has not only demonstrated that emotions are perceived categorically, but that this perception can be manipulated. The purpose of this study was to investigate, via two separate experiments using adaptation and multisensory techniques, whether the perception of body emotion expressions also shows categorical effects and plasticity. We used an approach developed for studies investigating both face and voice emotion perception and created novel morphed affective body stimuli, which varied in small incremental steps between emotions. Participants were instructed to perform an emotion categorisation of these morphed bodies after adaptation to bodies conveying different expressions (Experiment 1), or while simultaneously hearing affective voices (Experiment 2). We show that not only is body expression perceived categorically, but that both adaptation to affective body expressions and concurrent presentation of vocal affective information can shift the categorical boundary between body expressions, specifically for the angry body expressions. Overall, our findings provide significant new insights into emotional body categorisation, which may prove important in gaining a deeper understanding of body expression perception in everyday social situations.

Introduction

In daily life, we continuously interact with signals conveying emotion from multiple modalities, including the face, voice and body. However, to date the body has been the 'forgotten cousin' of the face and to a lesser extent, the voice, even though evidence confirms that we are well able to recognise affective information from bodily gestures and postures (de Gelder, 2009) and that there is an underlying neurobiological processing of body expression (de Gelder, 2006; Peelen & Downing, 2005, 2007; de Gelder et al., 2014). A number of studies utilising face and voice stimuli morphed between emotions (Laukka, 2005; Bestelmeyer et al., 2010; de Gelder & Vroomen, 2000) have shown that affective signals are divided by a distinct categorical boundary. However, categorical perception of body expressions has not yet been investigated. Furthermore, whether categorical perception can be shifted—as has been demonstrated for both faces and voices—is unknown.

✉ Beatrice de Gelder
b.degelder@maastrichtuniversity.nl

[1] Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, 6229 EV Maastricht, The Netherlands

Here, we investigated the perceptual representation of body affect, and in particular, the plasticity of the categorical boundary between different emotion categories. Indeed, a number of studies have shown that such boundaries can be manipulated using two different techniques. The first of these is adaptation, the process by which repeated exposure to a certain stimulus causes loss of sensitivity to those stimulus properties and a shifting of perception towards opposite features of the adapting stimulus (Grill-Spector et al., 1999). Adaptation paradigms have allowed us to gain significant insight as to how various sensory signals are organised and represented at the neural level: specifically, a number of studies have used this technique to uncover groups of neural populations that are tuned to specific stimulus attributes (Grill-Spector et al., 1999). Focussing on higher-order visual adaptation (specifically, studies using face stimuli), an early study by Webster and MacLin (1999) showed that continued adaptation to distorted faces (i.e., faces expanded along the horizontal and vertical dimensions) caused the dimensions of subsequently viewed faces to appear altered in the opposite direction (i.e., faces appeared contracted). Adaptation effects have now been extended to socially important cues such as facial identity (Leopold et al., 2001), attractiveness (Rhodes et al., 2003), expression (Fox & Barton, 2007;

Published online: 02 January 2019

Webster & MacLin, 1999), and gender and ethnicity (Webster & MacLeod, 2011).

High level auditory adaptation (specifically, that involving voice perception) has not received much attention. However, more recent evidence does suggest that aftereffects, equivalent to those seen for faces, can emerge when using vocal stimuli; for example, within the perception of gender (Schweinberger et al., 2008), identity (Belin & Zatorre, 2003) and affect (Bestelmeyer et al., 2010). With regards to voice affect, Bestelmeyer et al. showed that adaptation to angry vocalisations caused voices, morphed on an angry-fearful continuum, to be perceived as more fearful, and vice versa. Furthermore, a second experiment using caricatures showed that aftereffects were not exclusively due to low-level adaptation, but rather appeared to depend on a higher-level perception of the affective category of the adaptor.

Another influence on these perceptual boundaries is concurrent information from another source of affective signals, simultaneously paired with the unimodal stimuli under question. This work has mainly concentrated on face-voice combinations. For example, focussing on work using morphed stimuli, in an early study de Gelder and Vroomen (2000) presented participants with faces morphed between happiness and sadness, either uncoupled with a voice or simultaneously presented with either a happy or sad voice, and asked them to indicate whether they thought the person was happy or sad. Here, participants appeared to combine both sources of information, with categorisation of each face (apart from those completely congruent) shifted in the direction of the simultaneously presented voice, even when instructed to ignore it. Similarly, Ethofer et al. (2006) presented participants with either faces alone, voices alone or faces and voices together after which participants gave an emotional valence rating for each stimulus. Visual stimuli were from a morphed continuum of photographs, ranging from either happiness to neutral, or neutral to fear, and auditory stimuli were sentences spoken in either happy or fearful prosody. Participants rated fearful and neutral facial expressions as more fearful when presented concurrently with a fearfully spoken sentence.

Recently, these multisensory effects have been generalised to a broader domain by investigating affective cross-modal influences in whole-body expressions. Although research is to date limited, there is evidence that auditory information interacts with and influences the information we perceive from the body. For example, in a study by Van den Stock et al. (2008) participants viewed video clips of happy or fearful body language, simultaneously presented with either congruent or incongruent non-verbal human or animal vocalizations, or without auditory information. The results indicated that recognition of body language was biased towards the emotion expressed by the concurrently presented auditory information.

To explore categorical perception of body emotion, and the plasticity of the perceptual categories, we created a novel set of body stimuli, morphed between emotions. In Experiment 1, we studied adaptation effects: specifically, we hypothesised that if adaptation occurred, the baseline categorisation curve would be shifted in the opposite direction to the body emotion participants were adapted to. In Experiment 2, we investigated how simultaneously presented auditory information could affect the body morph categorisation curve. Here, we hypothesised that the baseline categorisation curve would shift towards the affective auditory information heard by the participant.

General materials and methods

Participants

Participants for both experiments were recruited online through Maastricht University. For Experiment 1 we recruited 25 participants (11 males, age range 18–36 years, mean ± S.D = 22.12 ± 2.60 years) and for Experiment 2 we recruited a new sample of 24 participants (12 males, age range 19–36 years, mean ± SD = 22.83 ± 3.58 years). The experiments were approved by the local ethical committee, and written informed consent was obtained before participation. Participants received monetary compensation in the form of a five Euro voucher.

Stimuli

Both Experiments 1 and 2 utilised emotional body morph continua based on independently validated body postures contained in the Bodily Expressive Action Stimulus Test ('BEAST'; de Gelder & Van den Stock, 2011). Specifically, 'Happy' and 'Angry' body postures from one male and one female were transferred to avatars using 3D Studio Max (Autodesk). The same software was used to morph between the two expressions to create one male and one female morphed continua, each consisting of 7 body postures. Particularly, we created an animation with the starting pose of the continuum as the first frame, and the ending pose as the last frame. The software then automatically calculated the transition in-between, and this transition was divided into the necessary number of steps and then rendered to images. In all stimuli, the faces were blurred so that the only source of affective information was the body posture. Examples of stimuli are shown in Fig. 1.

Statistical analyses

In both experiments, data for each participant were averaged as a function of each body-morphing step for each

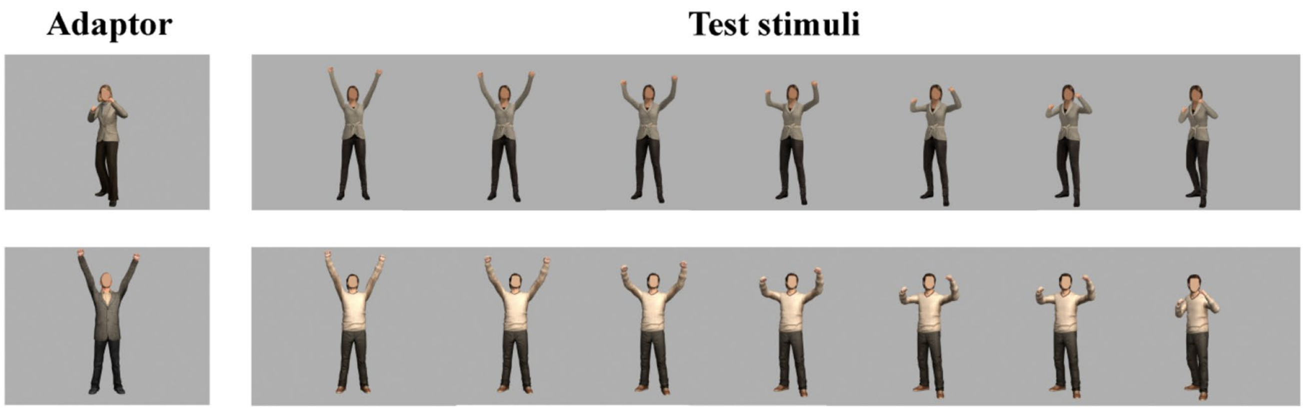

Fig. 1 Illustration of morphed affective body stimuli used in Experiments 1 and 2. In Experiment 1, participants were adapted to angry/happy and male/female bodies, and then tested on bodies drawn from a morphed affective continuum of another, within-gender identity. In Experiment 2, participants were tested on morphed affective bodies paired with affective vocal bursts

experimental condition (Experiment 1: body morph 1–7 with no adaptation, adaptation to anger, adaptation to happiness; Experiment 2: body morph 1–7 paired with a neutral voice, happy voice, and angry voice). A psychophysical curve based on the logistic function was then fitted for each experimental condition, for each participant, using Matlab R2012b by way of a multistep adaptive procedure. We, furthermore, computed for each of the individually fitted psychophysical curves, the point of subjective equality (PSE), which is the point along the morphing continuum that is the most ambiguous perceptually (yielding 50% of angry/happy responses). The mean group PSE for each condition, for both experiments, is illustrated in Fig. 2a, b with a cross on all the group average psychophysical curves. To determine differences between conditions within each experiment, we investigated whether there was a statistically significant difference between the PSEs of the fitted psychophysical curves for each condition by inputting these values into a one factor ANOVA. Greenhouse-Geisser corrected values are reported for ANOVAs in both experiments.

Experiment 1: Adaptation to body emotion

Methods

Stimuli

Male and female full body morph continua were used as test stimuli, whereas ‘Happy’ and ‘Angry’ avatars were used as adaptors. Adaptors were always the same gender as, but a different identity from the test stimuli. In total, we utilised 18 body stimuli: 4 adaptors (2 genders × 2 emotions) and 14 test stimuli (2 genders × 7 morphs).

Procedure

The experiment consisted of two main parts: a baseline emotion categorisation task without prior adaptation, and second the adaptation tasks. Both baseline and the adaptation task required a two-alternative forced choice (2-AFC) judgement of whether the body was Angry or Happy by means of a button press. The baseline task consisted of two blocks of trials, one for each gender and was always administered first. The body of each gender at each of the seven morph steps was repeated five times, leading to 35 trials per gender block. Within each block bodies were presented randomly, each for 800 ms, with an inter-stimulus interval of 2000 ms. Following the baseline task, we presented participants with the adaptation task. The trial structure of the adaptation task consisted of one adapting body presented four times in succession followed by a test stimulus. Each of the adapting and test bodies was presented for 800 ms, with a 200 ms inter-stimulus interval. The inter-trial interval was 5000 ms during which the participants judged the expression of the test stimulus. They were instructed to respond as soon as the test stimulus presentation was over. In total, there were four adaptation blocks (2 emotion × 2 gender) and each of the seven test stimuli per gender was repeated five times resulting in a total of 35 trials per block. Participants were always tested on a different identity than the one they were adapted to. The order of all baseline and adaptation blocks was counterbalanced; the baseline task by gender and the adaptation blocks by both gender and the adaptor emotion. All picture stimuli were presented at 720 × 576 pixels, using Presentation software (Neurobehavioral Systems, Inc.) running on a PC.

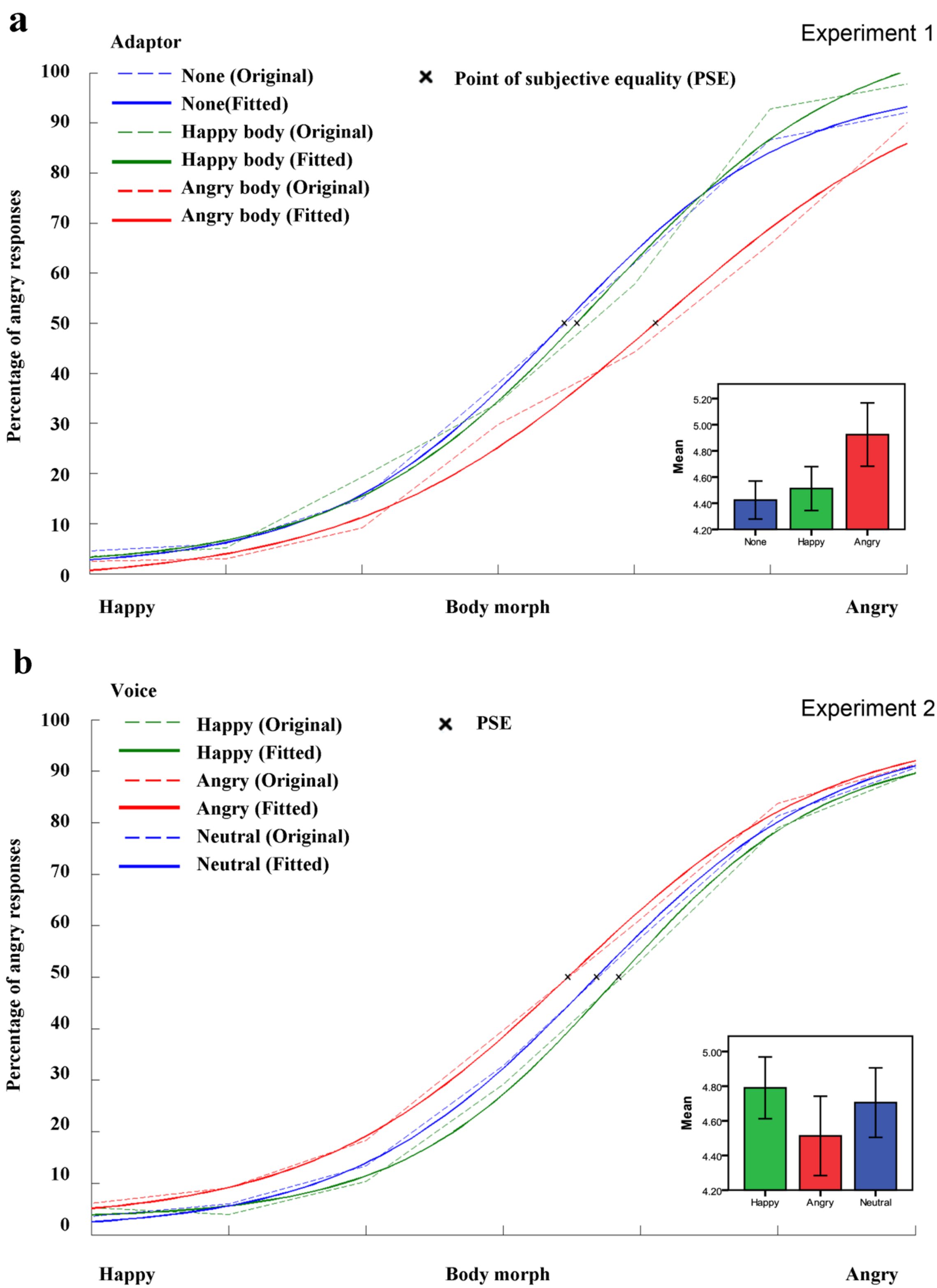

Fig. 2 **a** Original and fitted functions for each adaptation condition (blue = no adaptation; green = adaptation to happy bodies; red = adaptation to angry bodies). Percentage of 'angry' responses was plotted as a function of the happy-angry body morph continuum. The point of subjective equality (PSE) is shown as an '*x*' on each of the fitted functions. Inset figure displays the mean and SE of the PSE values; **b** Original and fitted functions for each audiovisual condition (blue = neutral voice; green = happy voice; red = angry voice). Again, the percentage of 'angry' responses was plotted as a function of the happy-angry body morph continuum, and the PSE is shown as an '*x*' on each of the fitted functions. Inset panels display the mean and SE of the PSE values

Results

The data from one participant were discarded due to the fact that we were unable to establish a PSE for their data for one condition. Figure 2a illustrates the group-averaged original categorisation curves for each adaptation condition and fitted function for each of these curves. First, we observed that in the 'unimodal' or baseline condition, a categorisation curve similar to those previously seen for morphed faces and voices emerged. Here, the proportion of angry responses appeared as a function of the proportion of 'angry' information in the body, with bodies centred around the ambiguous portion of the continuum being randomly categorised. The one factor ANOVA revealed a significant effect of adaptation to affective bodies ($F(1.52, 34.99) = 4.51$, $p < 0.03$). Exploring the main effects with Bonferroni correction showed that the PSE as a result of adaptation to angry bodies was significantly larger (i.e. more happy) than in the baseline condition ($p < 0.025$); however, adaptation to happy bodies did not elicit a significant shift in categorisation ($p = 0.47$).

Experiment 2: Simultaneous body-voice perception

Methods

Stimuli

In addition to the male and female body-morph continua, we also utilised auditory stimuli consisting of meaningless human vocalizations (i.e., "ah") expressing happy, angry or neutral affect, from one male and one female speaker, previously validated as part of another study (Stienen et al., 2011). The duration of each of the four bursts was 400 ms.

Procedure

This experiment consisted of 20 blocks (10 using female body morphs, 10 using male body morphs; blocks alternating between gender). In all blocks, each of the 7 morphs were paired with each of the three sounds (happy voice, angry voice, pure tone) resulting in a total of 21 stimuli. Every audiovisual stimulus was presented for 400 ms, with a 2000 ms inter-stimulus interval, and stimuli were randomised within each block. As in the previous experiment, participants performed a 2-AFC emotion categorisation task on the body, and were instructed to respond as quickly and as accurately as possible. Furthermore, they were explicitly instructed to ignore the simultaneously presented voice and base their decision only on the body affect. Participants could respond both when the stimulus was playing, and in the inter-stimulus interval. General experimental set-up (e.g. software etc.) was identical to Experiment 1.

Results

Figure 2b illustrates the group-averaged original categorisation curves for each audiovisual condition, and the fitted function for each of these curves. The one factor ANOVA revealed a significant effect of affective voices ($F(1.65, 37.88) = 5.53$, $p < 0.02$). Exploring the main effects with Bonferroni correction showed that the PSE in the angry voice condition was significantly smaller (i.e. more angry) than in the baseline condition ($p < 0.025$); however, pairing bodies with happy voices did not elicit a significant shift in categorisation ($p = 0.29$).

Discussion

Here, we investigated, via two separate experiments, how the perceptual boundary between different body emotions could be altered by either prior or concurrent presentation of affective signals. In Experiment 1, we observed a clear categorisation of body expression; furthermore, participants perceived an ambiguous body as significantly happier when they were adapted to an angry body. Adaptation effects have been demonstrated for both affective faces and voices; however, in this experiment we demonstrate for the first time that adaptation to affect in the body can influence the perception of a subsequently presented body expression. Interestingly, the observed effect appeared to be specific for the emotion considered, in that only angry bodies exerted a significantly adaptive effect. One suggestion for this result is that these body expressions have different functions—particularly, their implementations in action programmes—and underlying neural bases. Possibly different emotions exert differing adaptive effects, and it may ultimately be in our interest when we are interacting with our environment that we are less adaptable to certain emotions.

The fact that we can observe 'unimodal' (i.e. body to body) adaptation effects prompts the question as to whether we may be able to observe similar 'crossmodal' (e.g., voice/face to body) adaptation effects, as have previously been observed in studies using both faces and voices as test stimuli (Pye & Bestelmeyer, 2015; Skuk & Schweinberger, 2013). A recent study showed that adaptation to gender-specific faces could modulate the perception of a subsequently presented, androgynous body (i.e., after exposure to a female face, the androgynous body appeared as more male and vice versa; Palumbo et al., 2015). However, so far it was not clear whether such effects extend to body and voice emotion perception. Our novel set of morphed stimuli combined with adaptation techniques may provide evidence as to how bodies and voices are integrated at a more direct neural level, as has previously been done for faces and voices (Watson et al., 2014). Furthermore, morphed body continua may not only provide understanding as to how affect is integrated in healthy populations, but also allow for fine grained assessment of emotion recognition deficits in conditions such as autism and schizophrenia, previously established within other modalities (Poljak et al., 2012; Bediou et al., 2012; de Gelder et al., 2005).

In Experiment 2, we observed again that there is a clear categorisation of body expression and additionally showed that concurrently presented vocal affective information shifted the emotion perception of an ambiguous body. As in Experiment 1 the effects are different for angry and happy expressions. Here, it also seems that different expressions or emotions, concurrently presented affective vocal information—elicit different effects on body emotion categorisation. Our findings of emotion specificity are in line with a previous study (Ethofer et al., 2006) which found that when faces were paired with both happy and fearful vocalisations, it was only the fearful voices pairing that resulted in a change in emotion perception. These authors suggest this may be due to the fact that signals indicating danger or threat are arguably more biologically relevant than signals conveying positive emotions. It may similarly be the case that here angry vocalisations were inherently more salient, particularly considering the relatively short duration of the stimuli, and thus provoked a stronger shift in visual emotion categorisation. Still, at this stage and given the paucity of studies available, direct comparisons of emotions may be misleading. For example, it makes sense to assume that it is advantageous for the organism not to adapt to anger, which makes our finding in Experiment1 counterintuitive. However, to conclude this would mean to understand better how emotional signals play their role in naturalistic environments and, very importantly, what the optimal signal strength is that is needed for behavioral impact. More research is needed to create experimental conditions where the different emotion categories as well as the different sensory stimuli are functionally calibrated in terms of their optimal message strength. This is a major task for future studies of crossmodal affective processes.

In this experiment, we extend on previous research using morphed stimuli, for example faces (de Gelder & Vroomen, 2000; Ethofer et al., 2006) and voices (Bestelmeyer et al., 2010), and further develop the findings of Van den Stock et al. (2008) in that we used a more fine-grained set of stimuli (i.e., morphed affective bodies as opposed to only bodies that clearly expressed one emotion or the other) and that allowed us to determine shifts at an ambiguous point of body emotion categorisation. Furthermore, participants integrated vocal information even when instructed to attend to only the body. Previous behavioural (de Gelder & Vroomen, 2000; Collignon et al., 2008; Vroomen et al., 2001) and neural (Pourtois et al., 2000; Jessen & Kotz, 2011) evidence suggests that integration of face-voice emotion seems unaffected by attentional resources and is integrated at an early stage of processing; our findings appear to support this line of thinking in the domain of body-voice processing also. This behavioural research prompts the question as to the manner by which body and voice information is combined in the brain. This is under-researched; however, two recent studies using both electrophysiological (Jessen et al., 2012) and imaging (Jessen & Kotz, 2015) techniques suggests that bodies and voices are effectively integrated very early in processing, and that multisensory body-voice emotion perception has a distinct neural underpinning. The results of the current study are consistent with earlier studies showing that emotional signals merge. Recognition of emotional body expressions is influenced by the face (Kret, Stekelenburg, Roelofs & de Gelder, 2013, Experiment 3) and by the social emotional scene (Kret, Roelofs, Stekelenburg & de Gelder, 2013, Experiment 3) and emotionally incongruent face body compounds trigger an early ERP signal (Meeren, H. K. M., van Heijnsbergen, C., & de Gelder, B. 2005).

In conclusion, this research presents a number of novel theoretical advances. We show first that body emotion is perceived categorically, as has been shown previously for faces and voices; second, that these categories can be modified through visual adaptation; and third that these categories can be modified by auditory information provided by emotion in the voice. These results, coupled with future similar research within the area of body processing would no-doubt help the construction of a more detailed model of body perception, and allow comparisons between the processes underlying affect perception in multiple modalities.

Acknowledgements This work was supported by the European Research Council. The funding of the European Research Council was under the European Union's Seventh Framework Programme (FP7/2007–2013)/ERC grant agreement number 295673. We would also like to thank Dr. Julien Rouger for his helpful comments on this manuscript.

Compliance with ethical standards

Ethics statement All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

References

Bediou, B., Brunelin, J., d'Amato, T., Fecteau, S., Saoud, M., Hénaff, M., & Krolak-Salmon, P. (2012). A comparison of facial emotion processing in neurological and psychiatric conditions. *Frontiers in Psychology, 3*, 98.

Belin, P., & Zatorre, R. J. (2003). Adaptation to speaker's voice in right anterior temporal lobe. *Neuroreport, 14*, 2105–2109.

Bestelmeyer, P. E. G., Rouger, J., DeBruine, L. M., & Belin, P. (2010). Auditory adaptation in vocal affect perception. *Cognition, 117*, 217–223.

Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., & Lepore, F. (2008). Audio-visual integration of emotion expression. *Brain Research, 25*, 126–135.

de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience, 7*, 242–249.

de Gelder, B. (2009). Why bodies? Twelve reasons for including bodily expressions in affective neuroscience. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*, 3475–3484.

de Gelder, B., de Borst, A. W., & Watson, R. (2014). The perception of emotion in body expressions. *Wiley Interdisciplinary Reviews: Cognitive Science*. https://doi.org/10.1002/wcs.1335.

de Gelder, B., & van den Stock, J. (2011). The bodily expressive action stimulus test (BEAST). Construction and validation of a stimulus basis for measuring perception of whole body expression of emotions. *Frontiers in Psychology, 2*, 181.

de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion, 14*, 289–311.

de Gelder, B., Vroomen, J., de Jong, S. J., Masthoff, E. D., Trompenaars, F. J., & Hodiamont, P. (2005). Multisensory integration of emotional faces and voices in schizophrenics. *Schziphrenia Research, 72*, 195–203.

Ethofer, T., Anders, S., Erb, M., Droll, C., Royen, L., Saur, R., Reiterer, S., Grodd, W., & Wildgruber, D. (2006). Impact of voice on emotional judgment of faces: An event-related fMRI study. *Human Brain Mapping, 27*, 707–714.

Fox, C. J., & Barton, J. J. S. (2007). What is adapted in face adaptation? The neural representation of expression in the human visual system. *Brain Research, 1127*, 80–89.

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*, 187–203.

Jessen, J., & Kotz, S. (2015). Affect differentially modulates brain activation in uni- and multisensory body-voice perception. *Neuropsychologia, 66*, 134–143.

Jessen, S., & Kotz, S. A. (2011). The temporal dynamics of processing emotions from vocal, facial, and bodily expressions. *Neuroimage, 58*, 665–674.

Jessen, S., Obleser, J., & Kotz, S. A. (2012). How bodies and voices interact in early emotion perception. *PLoS One, 7*, e36070.

Kret, M. E., Roelofs, K., Stekelenburg, J., & de Gelder, B. (2013). Emotional signals from faces, bodies and scenes influence observers' face expressions, fixations and pupil-size. *Frontiers in Human Neuroscience, 7*, 810. https://doi.org/10.3389/fnhum.2013.00810.

Kret, M. E., Stekelenburg, J. J., Roelofs, K., & de Gelder, B. (2013). Perception of face and body expressions using electromyography, pupillometry and gaze measures. *Frontiers in Psychology*, *4*, 28. https://doi.org/10.3389/fpsyg.2013.00028.

Laukka (2005). Categorical perception of vocal expressions. *Emotion, 5*, 277–295.

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*, 89–94.

Meeren, H. K. M., van Heijnsbergen, C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. Proceedings of the National Academy of Sciences of the USA, 102, 16518–16523.

Palumbo, R., D'Ascenzo, S., & Tommasi, L. (2015). Cross-category adaptation: Exposure to faces produces gender aftereffects in body perception. *Psychological Research Psychologische Forschung, 79*, 380–388.

Peelen, M. V., & Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *Journal of neurophysiology, 93*, 603–608.

Peelen, M. V., & Downing, P. E. (2007). The neural basis of visual body perception. *Nature Reviews Neuroscience, 8*, 636–648.

Poljak, E., Poljak, E., & Wagemans, J. (2012). Reduced accuracy and sensitivity in the perception of emotional facial expressions in individuals with high autism spectrum traits. *Autism, 17*, 668–680.

Pourtois, G., de Gelder, B., Vroomen, J., Rossion, B., & Crommelinck, M. (2000). The time-course of intermodal binding between seeing and hearing affective information. *Neuroreport, 11*, 1329–1333.

Pye, A., & Bestelmeyer, P. (2015). Evidence for a supra-modal representation of emotion from cross-modal adaptation. *Cognition, 134*, 245–251.

Rhodes, G., Jeffery, L., Watson, T. L., Clifford, C. W. G., & Nakayama, K. (2003). Fitting the mind to the world: Face adaptation and attractiveness aftereffects. *Psychological Science, 14*, 558–566.

Schweinberger, S. R., Casper, C., Hauthal, N., Kaufmann, J. M., Kawahara, H., Kloth, N., et al. (2008). Auditory adaptation in voice perception. *Current Biology, 18*, 684–688.

Skuk, V. G., & Schweinberger, S. R. (2013). Adaptation aftereffects in vocal emotion perception elicited by expressive faces and voices. *PLoS One, 8*, e81691.

Stienen, B. M. C., Tanaka, A., & de Gelder, B. (2011). Emotional voice and emotional body postures influence each other independently of visual awareness. *PLoS ONE, 6*, e25517.

Van den Stock, J., Grèzes, J., & de Gelder, B. (2008). Human and animal sounds influence recognition of body language. *Brain Research, 1242*, 185–190.

Vroomen, J., Driver, J., & de Gelder, B. (2001). Is cross-modal integration of emotional expressions independent of attentional resources? Cognitive. *Affective and Behavioural Neurosciences, 1*, 382–387.

Watson, R., Latinus, M., Noguchi, T., Garrod, O., Crabbe, F., & Belin, P. (2014). Crossmodal adaptation in right posterior superior temporal sulcus during face-voice emotional integration. *Journal of Neuroscience, 34*, 6813–6821.

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature, 428*, 557–561.

Webster, M. A., & MacLin, O. H. (1999). Figural aftereffects in the perception of faces. *Psychonomic Bulletin & Review, 6*, 647–653.

Webster, M. A., & MacLeod, D. I. A. (2011). Visual adaptation and face perception. *Phil. Trans. R. Soc. B*, *366*, 1702–1725.