Observers perceive the Duchenne marker as signaling only intensity for sad expressions, not genuine emotion

Elizabeth J. Miller[1], Eva G. Krumhuber[2], and Amy Dawel[1]*

Revised submission to Emotion March 2020.

- Elizabeth J. Miller: elizabeth.miller@anu.edu.au; ORCID iD: 0000-0003-2572-6134
- Eva Krumhuber: e.krumhuber@ucl.ac.uk; ORCID iD: 0000-0003-1894-2517
- Amy Dawel: amy.dawel@anu.edu.au; ORCID iD: 0000-0001-6668-3121

[1]Research School of Psychology, The Australian National University, ACT 2600, Australia.

[2]Department of Experimental Psychology, University College London, 26 Bedford Way, London WC1H 0AP, United Kingdom.

*Corresponding author: Research School of Psychology (building 39), The Australian National University, Canberra, ACT 2600, Australia. tel. +61 2 6125 4106 e-mail: amy.dawel@anu.edu.au

Prepublication version of manuscript as accepted to Emotion 2021.

Keywords: Facial expression; genuine; Duchenne; AU6; intensity.

Abstract

The Duchenne marker—crow's feet wrinkles at the corner of the eyes—has a reputation for signaling genuine positive emotion in smiles. Here, we test whether this facial action might be better conceptualized as a marker of emotional intensity, rather than genuineness per se, and examine its perceptual outcomes beyond smiling, in sad expressions. For smiles, we found ratings of emotional intensity (how happy a face is) were unable to fully account for the effect of Duchenne status (present vs absent) on ratings of emotion genuineness. The Duchenne marker made a unique direct contribution to the perceived genuineness of smiles, supporting its reputation for signaling genuine emotion in smiling. In contrast, across four experiments, we found Duchenne sad expressions were not rated as any more genuine nor sincere than non-Duchenne ones. The Duchenne marker did however make sad expressions look sadder and more negative, just like it made smiles look happier and more positive. Together, these findings argue the Duchenne marker has an important role in sad as well as smiling expressions, but is interpreted differently in sad expressions (contributions to intensity only) compared to smiles (emotion genuineness independently of intensity).

(abstract word count=191, of max=250)

Observers perceive the Duchenne marker as signaling only intensity for sad expressions, not genuine emotion

An influential idea in the facial expression literature is that crow's feet wrinkles around the eyes signal genuine happiness when coupled with a smile (Darwin, 1872/2009; Duchenne de Boulogne, 1862/1990; Ekman, Davidson, & Friesen, 1990). This facial action—known as the Duchenne marker (Ekman et al., 1990)—is caused by contraction of the outer portion of the muscles that circle the eyes, *orbicularis oculi pars lateralis*, and also involves "bagging" of the skin under the eyes (Figure 1). Although the Duchenne marker has traditionally been conceptualised as signalling genuine happiness in smiles (for meta-analysis, see Gunnery & Ruben, 2016), some researchers have argued it communicates the intensity of emotion—how happy a person is—rather than genuineness per se (Fridlund, 1994; Messinger, 2002; Messinger et al., 2012). There is also considerable evidence the Duchenne marker appears in expressions of sadness and grief (Darwin, 1872/2009; Keltner & Bonanno, 1997; Namba, Kagamihara, Miyatani, & Nakao, 2017), raising questions about what this marker communicates in non-smiling expressions. In the present series of experiments, we aimed to address two questions about how the Duchenne marker influences observer perceptions of emotion. First, is the contribution of the Duchenne marker to perceived genuineness better accounted for by perceived intensity? And second, what are the perceptual effects of the Duchenne marker for sad expressions?

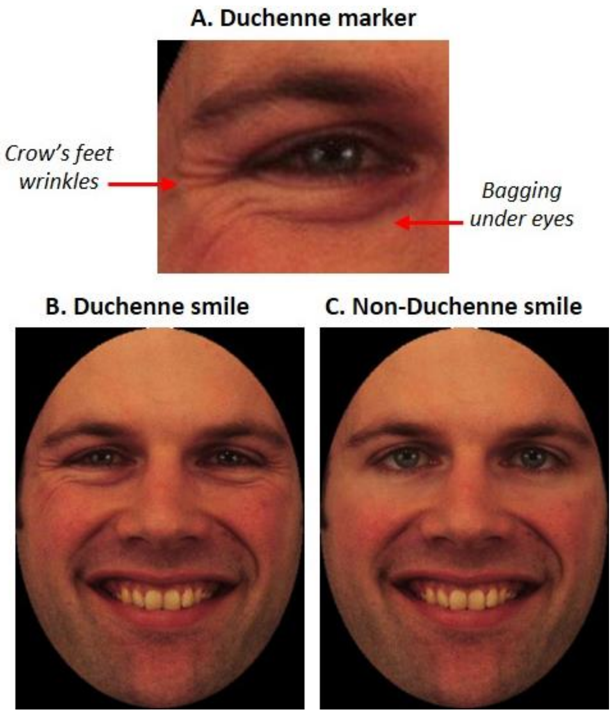

*Figure 1. (*A) shows a magnified view of the Duchenne marker, pointing out its key physical features. (B) and (C) show examples of Duchenne and non-Duchenne smile stimuli respectively (identity M12 from the KDEF, Lundqvist, Flykt, & Öhman, 1998a; http://kdef.se/home/using%20and%20publishing%20kdef%20and%20akdef.html gives permission for the KDEF images to be published in scientific journals).

Duchenne smiling

The idea that genuine happiness is signalled by contracting the orbicularis oculi in combination with the zygomaticus major (smile) muscle originated with the French anatomist Duchenne de Boulogne (1862/1990). Duchenne (1862/1990) claimed the inferior part of the orbicularis oculi muscle "does not obey the will; it is only brought into play by a true feeling, by an agreeable emotion" (p. 72). This idea, initially picked up by Darwin (1872/2009), was popularised by Paul Ekman and colleagues in the 1980s (Ekman et al., 1990; Ekman & Friesen, 1982a, 1988), sparking a substantial empirical literature on Duchenne smiling. This literature has two foci. One line of evidence focuses on the extent to which Duchenne smiling is associated with displayer self-reports of positive emotion, and occurs in contexts expected to elicit positive emotions. This body of evidence shows the frequency of Duchenne smiling correlates positively with how happy or amused people report feeling (Ekman et al., 1990; Harris & Alvarado, 2005; Matsumoto & Willingham, 2009) and is greater in contexts expected to elicit positive emotions compared to negative ones. For example, Duchenne smiles are more frequent in response to pleasant than unpleasant films (Ekman et al., 1990), wins versus losses (in Paralympians, Matsumoto & Willingham, 2009; and in children, Schneider & Unzer, 1992), hearing a joke versus experiencing pain (Harris & Alvarado, 2005), and when describing a happy event versus one which elicited anger (Hess & Bourgeois, 2010).

The second line of evidence, like the current study, focuses on perceiver interpretations of Duchenne smiles compared to smiles without the Duchenne marker (i.e., non-Duchenne smiles). Studies of Western participants unanimously show Duchenne smiles are perceived as more genuine and authentic than non-Duchenne ones (Quadflieg, Vermeulen, & Rossion, 2013; Song, Over, & Carpenter, 2016; Wang, Xu, Cui, Wang, & Ouyang, 2017). People also tend to attribute more positive characteristics to Duchenne than non-Duchenne smilers (for meta-analysis, see Gunnery & Ruben, 2016). For example, Duchenne smilers are rated as more likable (Frank, Ekman, & Friesen, 1993), extroverted and generous (Mehu, Little, & Dunbar, 2007), attractive and intelligent (Quadflieg et al., 2013), and prosocial (Song et al., 2016) than non-Duchenne smilers. Altogether then, the evidence seems compelling that Duchenne smiles communicate genuine happiness.

There is however a catch: Duchenne smiles are perceived not just as more genuine than non-Duchenne ones, but also as more intense (Gunnery, Hall, & Ruben, 2013; Krumhuber & Manstead, 2009; Quadflieg et al., 2013). For this reason, some researchers (Fridlund, 1994; Messinger, 2002; Messinger et al., 2012) have argued the Duchenne marker should be conceptualised as a marker of emotional intensity—in the context of smiles, feeling *more* happy or more positive emotion—rather than emotional genuineness per se. This argument implies that the effect of Duchenne status (present vs. absent) on perceived intensity can account for the apparent association between the Duchenne marker and perceived genuineness. The present study will directly investigate this hypothesis for the first time, using mediation analyses to test whether perceived intensity explains (mediates) the relationship between Duchenne status and perceived genuineness.

Duchenne expressions of sadness and grief

A second challenge to the idea that the Duchenne marker signals genuine happiness with any specificity is that it also appears in non-smiling expressions, most notably sad expressions (Darwin, 1872/2009; Keltner & Bonanno, 1997; Namba et al., 2017). Again, it was initially Darwin's observations that tied the contraction of the orbicularis oculi to crying and grief (Darwin, 1872/2009). Modern empirical work has followed-up with strong evidence that the Duchenne marker frequently appears in expressions of sadness and in contexts involving loss. In total, we located fourteen studies that found the Duchenne marker appeared in such expressions. Six studies found the Duchenne marker was produced in response to sad mood induction procedures (listening to gloomy music and recalling a sad experience; Namba et al., 2017; talking about a loved one who had recently died; Keltner & Bonanno, 1997; describing a negative experience where someone died; Lee & Beattie, 1998; watching distressing videos; Ekman et al., 1990; looking at sad images, Khan, Ward, & Ingleby, 2009; and during a monologue task; Papa & Bonanno, 2008). Five studies found the Duchenne marker was included in actors' posed expressions of sadness (Carroll & Russell, 1997; Grogorick, Albuquerque, Tauscher, Kassubeck, & Magnor, 2019; Mehu, Mortillaro, Bänziger, & Scherer, 2012; Scherer, Ellgring, Dieckmann, Unfried, & Mortillaro, 2019) or despair (Scherer & Ellgring, 2007). Two studies found the Duchenne marker in expressive responses to losing games (tennis; Aviezer et al., 2015; an achievement game; Schneider & Josephs, 1991). Finally, one study found the Duchenne marker appears in infant cry-faces (Mattson, Cohn, Mahoor, Gangi, & Messinger, 2013). Overall then, there is good evidence the Duchenne marker appears in expressions of sad emotion and during events involving loss.

Evidence regarding the perceptual outcomes of the Duchenne marker in such expressions is however scarce. To the best of our knowledge, there are only four relevant studies: one testing sad expressions displayed by "FaceGen" computer-generated adult faces (Malek et al., 2019) and three testing infant cry-faces (Dinehart et al., 2005; Messinger, 2002; Messinger et al., 2012). All four studies tested perceived emotional intensity, and found Duchenne sad expressions or cry-faces were perceived (rated) as more negative (Dinehart et al., 2005; Malek et al., 2019; Messinger et al., 2012) or more upset-distressed (Messinger, 2002) than non-Duchenne versions of the expressions. Importantly, the studies tested a range of different stimuli: original photographs of naturally-occurring Duchenne and non-Duchenne crying (Dinehart et al., 2005; Messinger et al., 2012); images of Duchenne crying that were created by digitally transplanting the Duchenne marker into non-Duchenne cry faces, so that the other information in the faces was the same (e.g., intensity of mouth activation; Dinehart et al., 2005; Messinger, 2002); and FaceGen images, in which the expressive actions of the face were manipulated using FACSGen software (Malek et al., 2019). Thus the evidence is strong and robust that the Duchenne marker makes sad expressions appear more intensely negative/upset. For perceptions of emotion genuineness however, only the Malek et al. (2019) study testing FaceGen images collected relevant data. Namely, ratings of sincerity. Malek et al. (2019) found Duchenne sad expressions were rated as more sincere, as well as more negative, than non-Duchenne ones—just as Duchenne smiles were rated as more sincere, as well as more positive, than non-Duchenne ones. This identical pattern of results for sadness and smiles was interpreted by Malek et al. (2019) as evidence that the Duchenne marker has the same communicative function in both expressions, communicating both sincerity and affective intensity. In relation to sad expressions, the main aim of the present study is to retest whether Duchenne status is associated with perceived genuineness using a different type of stimulus image. In particular, photographic images of 40 human face identities from a popular photographic database, the Karolinska Directed Emotional Faces (KDEF; Lundqvist et al., 1998a), in which we manipulated the presence/absence of the Duchenne marker using the digital transplantation procedure from studies of infant cry-faces (Dinehart et al., 2005; Messinger, 2002).

The present research

In sum, our series of four experiments aimed to address two important gaps in the literature. First, there has been no direct test of the idea that the Duchenne marker might be better conceptualised as a marker of emotional intensity, rather than emotional genuineness. Here, we collect ratings of perceived intensity and genuineness from the same participants, for Duchenne and non-Duchenne smiles, and test this idea using mediation analysis. If the Duchenne marker really is a marker of emotional intensity rather than genuineness, then rated intensity should fully explain (mediate) the relationship between Duchenne status (present vs absent) and rated genuineness.

Second, little is known about the role of the Duchenne marker in the perceived genuineness of sad expressions. The single study on this topic used computer-generated faces and found Duchenne sad expressions were rated as more sincere than non-Duchenne ones (Malek et al., 2019). Here, we re-tested the hypothesis that the Duchenne marker influences the perceived genuineness of sad expressions using photographs of real people from the Karolinska Directed Emotional Faces database (KDEF; Lundqvist et al., 1998a) and manipulating the presence/absence of the Duchenne marker using an image editing tool (e.g., as in Dinehart et al., 2005; Messinger, 2002). Based on Malek et al.'s (2019) findings, we hypothesised that Duchenne sad expressions would be rated as more genuine than sad expressions without this marker.

Experiment 1: Genuineness, emotional and physical intensity ratings

Experiment 1 used a series of rating tasks to investigate the perceived genuineness and intensity of Duchenne relative to non-Duchenne smiling and sad expressions. Genuineness ratings used Dawel et al.'s (2017) task, which include detailed instructions about what genuine and fake/acted/posed expressions are, making clear to participants precisely what we wanted them to rate. For intensity, past work has typically asked participants to rate one type of intensity, with minimal instruction (e.g., valence intensity, Malek et al., 2019; Messinger et al., 2012; or smile intensity, Quadflieg et al., 2013). A concern about collecting ratings of intensity in this way is that it may confound perceptions of emotional intensity (how strongly a face communicates an emotion) with the perceived physical intensity of facial configurations (how much facial muscles are moved away from their neutral, resting position). Because we were primarily interested in testing the role of emotional intensity (Fridlund, 1994; Messinger, 2002; Messinger et al., 2012) in our mediation analyses, we separated out these two potential types of intensity into separate rating tasks. First, for emotional intensity, we asked participants to rate how strongly each face was showing different emotions, focusing on ratings of happiness for smiles, and ratings of sadness for sad expressions. Then, for physical intensity, we asked participants to rate how much the emotional facial expressions differed physically from neutral.

Method

Participants

Data analysed were from 18 undergraduate students (67% female, 33% male; *M* age=19.2 years, *SD*=1.3, range 18-22) who received course credit for the single 2.5 hour test session. Data from two additional participants were excluded due to a technical problem (1) or excessive missing data (1, missing 20% of emotion rating trials because they manually skipped them). Inclusion/exclusion criteria for all four experiments required that participants: be Caucasian, to match the race of our face stimuli, and raised in Australia or other majority-Caucasian countries such as New Zealand, UK, Canada and USA (because there can be cultural differences in perceptions of emotion genuineness; Thibault, Levesque, Gosselin, & Hess, 2012); report normal or corrected-to-normal vision; and report not having a clinical or neurological disorder likely to impair face task performance (exclusion criteria included Autism Spectrum Disorder, ADD/ADHD, being currently medicated for a psychological disorder, or having a major neurological condition or injury, e.g., epilepsy, stroke). The four experiments were approved by the ANU Human Research Ethics Committee.

Face stimuli

Stimuli were Duchenne and non-Duchenne smiles, as well as Duchenne and original sad expressions. Figures 1 and 2 show examples of our smiling and sad expression stimuli respectively. We used the 40 different identities (20 males, 20 females) selected by Calvo & Fernández-Martín (2013) from the KDEF (Lundqvist et al., 1998a) because their happy expressions included the Duchenne marker (see also Calvo, Fernández-Martín, Recio, & Lundqvist, 2018, for evidence regarding the presence of the Duchenne marker in this stimulus set). The original KDEF photographs show expressions elicited by instructions to "evoke the emotion [to be] expressed, and – while maintaining a way of expressing the emotion that [feels] natural to [you] – try and make the expression strong and clear" (Lundqvist, Flykt, & Öhman, 1998b). Calvo and Fernández-Martín (2013) resized the KDEF photographs so that the eyes were the same size and located at the same position in all of the images of a given person. Then, they cropped the images to a standard oval shape that isolated the internal region of the faces, removing hair and other background information. All of our stimuli derived from these size-standardised, oval-cropped versions of the KDEF photographs.

Duchenne expressions. The Duchenne smiles were Calvo & Fernández-Martíns' (2013) oval-cropped versions of the original KDEF happy expressions. We created Duchenne sad expressions (Figure 2A) by cutting the Duchenne eyes out of a person's happy expression image and transplanting them into the original sad image of that same person, using Photoshop CC Version 2017.1.0 (Knoll & Knoll, 2018). We were careful to transplant in only those bits around the eyes that were part of the Duchenne marker (i.e., crow's feet wrinkles, and bagging under the eyes; not editing the original eyebrows in any way). The "donor skin" from the transplanted eye region was then blended into the surrounding skin of the receiver face.

Comparison expressions. Non-Duchenne smiles were created by cutting the eyes out of a person's neutral expression image and transplanting them into the original happy image of that same person, using Photoshop CC Version 2017.1.0 (Knoll & Knoll, 2018). The comparison sad expressions used in Experiment 1 were Calvo and Fernández-Martíns' (2013) oval-cropped versions of the original KDEF photographs (i.e., original sad expressions, as shown in Figure 2B).

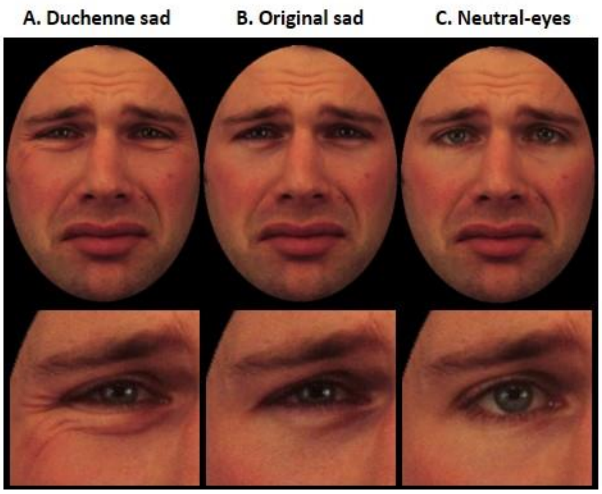

Figure 2. Top row shows examples of our Duchenne, original and neutral-eyes sad expression stimuli. Bottom row shows a magnified view of how the eye region differed across conditions. The stimulus identity is M12 from the KDEF (Lundqvist, Flykt, & Öhman, 1998a; http://kdef.se/home/using%20and%20publishing%20kdef%20and%20akdef.html gives permission for the KDEF images to be published in scientific journals).

Experiment presentation

For all experiments, the face rating tasks were presented using PsychoPy Version 1.90.1 (Peirce, 2007, 2008) on a 27-inch Apple iMac, at 5120 x 2880 pixels and 60hZ. Faces were displayed at 9° v.a. high x 7° v.a. wide (9.5cm x 7cm at viewing distance of ~60cm). This viewing angle was chosen to reflect the approximate size that people most frequently see facial expressions at in real life (based on average Caucasian face size of 18.5cm x 13.8cm, Weinberg et al., 2016, viewed at typical conversational distance of ~1 m for majority-Caucasian countries, Sorokowska et al., 2017).

Experimental procedure: Face rating tasks

General procedure for all four experiments. In all of the face tasks, each trial presented a single face image until response. Each image was shown once per task, in a different random order for each participant, with the condition that two images of the same person not be shown in direct succession. In each face task, participants rated all of the face images before they moved onto the next task. After each participant had completed all of the rating tasks in their experiment, we asked if they had noticed anything unusual about the faces they just saw. No participant in any experiment reported being aware of the eye transplants we had performed on the face images. The experimental sessions ended with a brief demographic questionnaire (age, sex, and questions that verified participants met inclusion/exclusion criteria).

The Experiment 1 tasks are described briefly below, in the order they were completed. Full instructions are presented in Supplement S2. Experiment 1 also tested images of other emotional expressions (anger, disgust, fear, surprise) for exploratory purposes. Results for these stimuli are reported in Supplement S3.

Emotional intensity ratings. Participants rated how strongly each face showed each of six emotions—anger, disgust, fear, happiness, sadness and surprise—from 0/1 *none/weak* to 10 *strong*.

Genuineness ratings. Participants rated the genuineness of the emotion being expressed by each face from –7 *completely fake* through 0 *don't know* to +7 *completely genuine*, following Dawel et al.'s (2017) procedure. Instructions at the beginning of the task provided a detailed explanation of emotional genuineness, including:

> *"Sometimes facial expressions reflect genuinely felt emotion, but other times expressions are faked or posed (e.g., to be polite)…An example of a genuine expression is when somebody smiles and they really feel happy, like when they get a present or see something funny. An example of a faked expression is when somebody smiles for a school photo, without feeling any emotion… All the expressions in this task were photographed in the lab, but some of them are genuine and some are faked. In genuine expressions, emotions (including mixed emotions) were induced by exposing people to video clips, pictures, or smells, or by having people recall and 'relive' an emotional event. For example, some genuine happy expressions were photographed in response to funny videos… In faked expressions, people were simply instructed to act different emotions (including mixed emotions). For example, some people showing faked happy expressions were instructed to pose for a photo."*

Participants were also instructed to ignore the physical strength of expressions:

> *"For example, an expression of sadness may be very subtle but completely genuinely felt… On the other hand, an expression of sadness may be very strong but be completely faked/posed/acted."*

Physical intensity ratings. Participants rated the physical intensity of the expressions from 0/1 (*none/weak)* to 10 (*strong)*. Instructions at the beginning of the task defined physical intensity as the extent to which an emotional expression differed from a neutral one.

Data screening and analysis

To ensure genuineness ratings reflected the perceived genuineness of the intended emotion for each stimulus, we retained genuineness ratings only if the participant rated that stimulus as showing the intended emotion more strongly than any other emotion (12% of trials removed in Exp. 1). For example, we retained a participant's genuineness rating for a sad image only if they had rated that image as being more sad than angry, fearful etcetera. Note, for all four experiments, re-analysing our data with all ratings retained had minimal impact on results, and did not alter our key findings.

Our main analyses used mean ratings for individual stimuli, calculated by averaging the ratings for each stimulus across participants.[1] Paired-sample *t*-tests were used to compare mean ratings (e.g., of genuineness) for Duchenne versus comparison expressions (e.g., Duchenne vs non-Duchenne smiles, or Duchenne vs original sad expressions). The pairing in these *t*-tests relates to the identity of the stimulus face (i.e., ours is a within-subjects design, but where the "subjects" are the identities of the faces, not the participants).

Power

We tested N=18 raters in each experiment to err on the conservative side of DeBruine and Jones' (2018) work showing that N=15 raters are typically required to obtain stable mean ratings for individual stimuli. Because our analyses compared mean ratings for individual stimuli across conditions (rather than mean ratings for individual participants), the purpose of power analyses was to identify the number of stimulus identities (rather than the number of participants) required. To achieve power of 0.8 for a medium effect (Cohen's d=0.5) with α=.05 we required N=34 (calculated with G*Power, Faul, Erdfelder, Buchner, & Lang, 2009; Faul, Erdfelder, Lang,

[1] We conducted further analyses for each experiment using the mean ratings for each participant, in which we averaged the ratings for each participant across the stimuli. Results are detailed in Supplement S6 and largely replicate the pattern of findings presented herein. Critically, these analyses similarly showed no effect of Duchenne status on ratings of genuineness for sad expressions.

& Buchner, 2007). Thus our sample size of 40 face identities in Experiments 1, 2 and 4 exceeded the required power. Experiment 3 tested only 20 identities, but all Experiment 3 tasks were re-tested with 40 identities in Experiment 4, which produced the same results.

Results

Replication of typical findings for Duchenne compared to non-Duchenne smiles

Figure 3A presents results for smiles. As is typically found, Duchenne smiles were perceived (rated) as more genuine than non-Duchenne ones, $M_{Duchenne}$=2.17, $M_{non\text{-}Duchenne}$=0.89, $t(39)$=5.53, p<.001, d=.87, as well as more happy, $M_{Duchenne}$=7.19, $M_{non\text{-}Duchenne}$=6.72, $t(39)$=5.22, p<.001, d=.82, and more physically intense, $M_{Duchenne}$=5.69, $M_{non\text{-}Duchenne}$=5.53, $t(39)$=2.52, p=.016, d=.40.

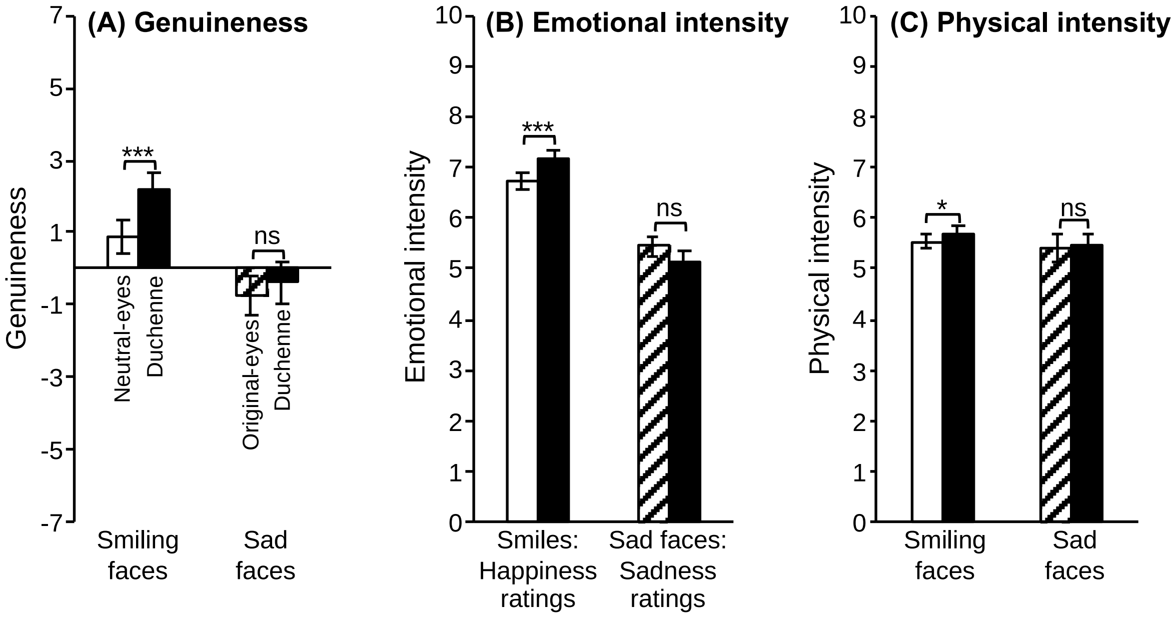

Figure 3. Experiment 1: mean ratings of (A) genuineness, (B) emotional intensity, and (C) physical intensity for smiling and sad faces. Mean ratings were derived by averaging the participant ratings for each stimulus first, and then averaging across all stimuli. Brackets above each pair of bars indicate results from paired samples *t*-tests. Error bars show 95% CIs for the mean difference from *t*-tests. *p <.05. ***p <.001. ns=non-significant.

Does the Duchenne marker signal emotional genuineness over and above intensity for smiles?

The above analyses provide clear evidence that our Duchenne smiles were perceived as more emotionally intense (i.e., more happy) as well as more genuine than the non-Duchenne ones, with both effects being similarly large (i.e., happiness ratings *d*=.82, genuineness ratings *d*=.87). These findings are potentially consistent with the idea that the Duchenne marker might be better conceptualised as a marker of emotional intensity (Fridlund, 1994; Messinger, 2002; Messinger et al., 2012) rather than genuineness per se. To test this hypothesis directly, we used mediation analyses to establish whether the effect of Duchenne status on perceived genuineness could be accounted for (was mediated) by its effect on perceived intensity. Figure 4 illustrates the results of our mediation model (*n* bootstraps=10,000; run using Hayes' PROCESS macro, (Hayes, 2018). Although the bootstrap analysis for the indirect effect produced a 95% CI that did not cross zero, *b* = .28, 95% CI = [.08, .61], which is indicative of mediation, the main finding was that the direct effect of Duchenne status on perceived genuineness remained strong and significant, *b* = 1.00, 95% CI = [.36, 1.64], *p* = .003. Sobel's test only approached significance (*z* = 1.88, *p* = .06), suggesting that mediation was partial. Thus, the Duchenne marker had a unique direct effect on how genuine smiles were perceived to be, over and above any contributions made via its effect on perceived intensity.

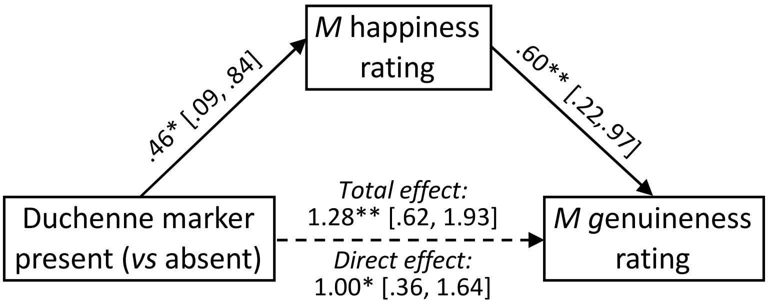

Figure 4. Mediation analysis showing that the direct association of Duchenne status with perceived genuineness remained strong and significant despite partial mediation via emotional intensity (happiness ratings). Thus perceived intensity was unable to fully account for perceptions of smile genuineness. Total effect = effect of X on Y without mediator. Direct effect = effect of X on Y given mediator. Path strength coefficients are unstandardized weights. *p<.05. **p<.01.

No effect of the Duchenne marker on perceptions of sadness

Figure 3B presents results for sadness. Unlike for smiles, Duchenne sad expressions were not perceived (rated) as more genuine nor intense than their comparison expressions. There were no significant differences between Duchenne and original sad expressions for ratings of genuineness, $M_{Duchenne}$=-0.39, $M_{original=}$-0.73, $t(39)=1.25$, p=.220, d=.20, sadness, $M_{Duchenne=}$5.15, $M_{original=}$5.45, $t(39)=1.68$, p=.102, d=.26, or physical intensity, $M_{Duchenne=}$5.45, $M_{original=}$5.40, $t(39)=0.44$, p=.661, d=.07.

Discussion

As expected, Experiment 1 showed Duchenne smiles were perceived as more genuine, happy and physically intense than non-Duchenne ones. Interestingly, the effects of Duchenne status on ratings of genuineness and happiness were similarly large in magnitude. However, mediation analyses clearly established the Duchenne marker makes a unique direct contribution to perceived genuineness that is not explained by the effects of Duchenne status on intensity. This finding provides strong evidence the conceptualisation of the Duchenne marker as a marker of smile genuineness, not just emotional intensity, is justified.

For sad expressions however, we found Duchenne status had no effect on perceptions of genuineness or intensity. That is, inserting the Duchenne marker into the sad expressions did not make them look any more genuine, nor sad, nor physically intense than the original sad expressions.

Experiment 2: Genuineness ratings for Duchenne vs neutral-eyes sad expressions

Our finding that Duchenne status had no effect on perceptions of sad expressions differs from that of Malek et al. (2019), who found the Duchenne marker made sad expressions look more sincere and more negative. One issue we identified with hindsight was that many of our "non-Duchenne" original sad expressions were not completely neutral around the eyes, as Malek et al.'s (2019) were. Instead, some of our stimuli showed hints of the Duchenne marker and/or a closely-related facial action, known as AU7. Recall, the Duchenne marker involves contraction of the outer *pars lateralis* portion of the orbicularis oculi, which is designated Action Unit 6 (AU6) by the Facial Action Coding System (FACS; Ekman, Friesen, & Hagar, 2002)—a system for precisely and objectively coding the possible independent actions of the face. In comparison to the Duchenne marker (AU6), AU7 involves contraction of the inner *pars palpebralis* portion of the orbicularis oculi, which also narrows the eye aperture and may cause under-eye bulging, but does not produce the full crow's feet wrinkles of the Duchenne marker (Ekman et al., 2002).

For this reason, we added neutral-eyes versions of the sad expressions into our follow-up experiments. These stimuli were created using the same procedure as we used for the non-Duchenne smiles. That is, by transplanting each person's eyes from their neutral expression image into their sad one. Because previous studies had already provided robust evidence Duchenne sad or cry expressions are perceived as more intense then neutral-eyes ones (Dinehart et al., 2005; Malek et al., 2019; Messinger, 2002; Messinger et al., 2012), Experiment 2 focused on ratings of genuineness.

Method

Participants

Data analysed were from 18 undergraduate students (78% female, 22% male; *M* age=19.4 years, *SD*=1.3, range 18-23) who had not participated in Experiment 1. Participants received course credit for the single 1 hour test session. All other participant details (e.g., inclusion criteria, ethics approval) were as for Experiment 1.

Face stimuli

Stimuli were the Duchenne and original sad expressions used in Experiment 1, plus neutral-eyes versions of the sad expressions. The neutral-eyes sad expressions were created using the same procedure we used for the non-Duchenne smiles in Experiment 1. That is, by cutting the eyes out of each person's neutral expression image and transplanting them into the original sad face for that same person, using Photoshop CC Version 2017.1.0 (Knoll & Knoll, 2018). Experiment 2 also tested images of other emotional expressions (anger, disgust, fear, surprise) for exploratory purposes. Results for these additional stimuli are reported in Supplement S3.

Genuineness ratings with forced-choice emotion labelling

For each face image, participants indicated what emotion they thought was being shown from six alternative labels—anger, disgust, fear, happy, sadness, or surprise—and then rated the genuineness of the emotion being expressed following the same instructions and using the same –7 to +7 scale as in Experiment 1.

Data screening and analysis

To ensure genuineness ratings reflected the perceived genuineness of the intended emotion for each stimulus, genuineness ratings were retained for analysis only if the participant labelled that stimulus as showing the intended emotion (18% removed). Paired-sample *t*-tests were used to compare mean ratings of genuineness for each combination of the three types of sad expressions (i.e., Duchenne vs original sadness, Duchenne vs neutral-eyes sadness, original vs neutral-eyes sadness). A Bonferroni corrected alpha level of .05/3=.017 was applied to adjust for the three comparisons.

Results and Discussion

Experiment 2 found that, even when compared to neutral-eyes sad expressions, people did not perceive Duchenne sadness as more genuine. Figure 3Bi shows there were no significant differences in ratings of genuineness amongst the three types of sad expressions: $M_{Duchenne=}$0.58, $M_{neutral\text{-}eyes}$= 0.21, $t(39)=1.15$, $p=.258$, $d=.18$; $M_{Duchenne=}$0.58, $M_{original}$ =0.22, $t(39)=1.09$, $p=.283$, $d=.17$; $M_{original=}$.22, $M_{neutral\text{-}eyes}$= .21, $t(39)=.03$, $p=.977$, $d=.004$. This set of results argue the Duchenne marker is not interpreted as communicating genuineness in sad expressions.

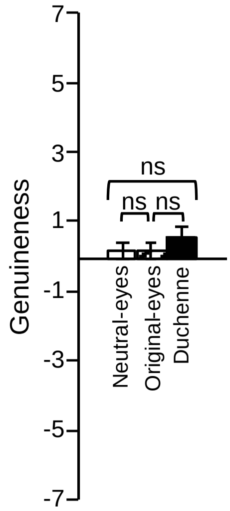

Figure 5. Experiment 2: mean ratings of genuineness for sad faces. Mean ratings were derived by averaging the participant ratings for each stimulus first, and then averaging across all stimuli. Brackets above each pair of bars indicate results from paired samples *t*-tests. Error bars show 95% CIs from repeated-measures ANOVA comparing all three conditions, calculated using the formula: t-crit*sqrt(MSE/N) (Loftus & Masson, 1994). Note, this ANOVA also showed no significant effect of Duchenne status across the three conditions, $F(2,78)=0.86$, $MSE=2.06$, $p=.426$. ns=non-significant.

Experiment 3: Malek et al.'s (2019) valence and sincerity rating tasks

Our finding that Duchenne status had no effect on the perceived genuineness of sad expressions even when the comparison expressions had neutral-eyes, as in Malek et al. (2019), remains at odds with Malek et al.'s (2019) results for sincerity. We next hypothesised that the difference between our findings and Malek et al.'s (2019) might be due to differences between our rating tasks and theirs. Our genuineness rating task asked participants to rate *emotional genuineness*, following detailed instructions about what we meant by this term (Dawel et al., 2017). In contrast, Malek et al. (2019) asked their participants rate *sincerity*, without the meaning of this term being specified (N. Malek, personal communication, September 12, 2018). Experiment 3 therefore collected ratings of sincerity, and also valence, using Malek et al.'s (2019) general procedure (see our Methods for details of scale extension), before collecting genuineness ratings using our own procedure. Importantly, participants completed Malek et al.'s (2019) valence and sincerity rating tasks first so that these ratings were not contaminated by our detailed instructions about genuineness. We also included smiles in Experiment 3 so that we could check our participants responded as expected to stimuli for which effects are already well-established (i.e., to make sure our participants were otherwise responding normally).

Method

Participants

Data analysed were 18 undergraduate students (61% female, 39% male; *M* age=20.7 years, *SD*=2.2, range 18-28) who had not participated in the previous experiments. Participants received course credit for the single 1 hour test session. All other participant details were as for Experiment 1.

Face stimuli

Stimuli were the Duchenne and neutral-eyes smiles and sad expressions from Experiments 1 and 2 respectively. We also included Calvo and Fernández-Martíns' (2013) oval-cropped versions of the original KDEF neutral expression images, as Malek et al. (2019) had included neutral expressions in their study. A key change from Experiments 1 and 2 was that we used only half of our original KDEF identities (10 males + 10 females=20 total identities). The reason for halving the number of identities was that we also tested computer-generated versions of the faces, created in FaceGen (Singular Inversions Inc., 2009). The 20 specific identities we used comprised the 10 males and 10 female identities that produced the best quality FaceGen images. Supplement 4 presents further details and results of the FaceGen image testing.

Experimental procedure

Valence ratings. Participants rated “How negative/positive do you find the stimulus?” from –7 *extremely negative* through 0 *neutral* to +7 *extremely positive*. We extended Malek et al.’s (2019) original 1 *very negative*—5 *very positive* scale because Likert-type scales function better with 7+ options (Preston & Colman, 2000). The addition of a neutral midpoint at 0 also enabled us to make clear interpretations about whether expressions were perceived as positively or negatively valenced (i.e., values >0=positive, values <0=negative).

Sincerity ratings. Participants rated “How sincere do you find the expression?” from 0 *not at all*—7 *very*. This scale was also extended from Malek et al.’s (2019) original 1 *not at all*—5 *very* scale because Likert-type scales function better with 7+ options (Preston & Colman, 2000).

Genuineness rating with forced-choice emotion labelling. This task was the same as in Experiment 2, except that “neutral” was added to the labelling options (because we included neutral expression images in testing).

Data screening and analysis

To ensure genuineness ratings reflected the perceived genuineness of the intended emotion for each stimulus, genuineness ratings were retained for analysis only if the participant labelled that stimulus as showing the intended emotion (22% removed). Paired-sample *t*-tests were used to compare mean ratings (e.g., of genuineness) for the Duchenne versus neutral-eyes versions of an expression (i.e., Duchenne vs neutral-eyes smiles, Duchenne vs neutral-eyes sadness).

Results

Figure 3 shows that results were as expected for smiles. Compared to neutral-eyes smiles, Duchenne smiles were rated as significantly more genuine, $M_{Duchenne=}$2.98, $M_{neutral\text{-}eyes}$=1.12, $t(19)=5.39$, $p<.001$, $d=1.21$, sincere, $M_{Duchenne=}$4.87, $M_{neutral\text{-}eyes=}$4.39, $t(19)=4.64$, $p=<.001$, $d=1.04$, and positive, $M_{Duchenne=}$4.31, $M_{neutral\text{-}eyes=}$3.84, $t(19)=4.66$, $p=<.001$, $d=1.04$.

For sadness however, the pattern of results was quite different. Again, we found no significant difference between Duchenne and neutral-eyes sad expressions for ratings of genuineness $M_{Duchenne}$=-0.85, $M_{neutral\text{-}eyes=}$-0.59, $t(19)=0.59$, $p=.562$, $d=.13$, and sincerity, $M_{Duchenne}$=3.38, $M_{neutral\text{-}eyes}$=3.24, $t(19)=1.17$, $p=.255$, $d=.26$. We did however find that Duchenne sad expressions were rated as more negative than neutral-eyes ones, $M_{Duchenne}$=-2.84, $M_{neutral\text{-}eyes}$=-2.58, $t(19)=3.18$, $p=.005$, $d=.71$.

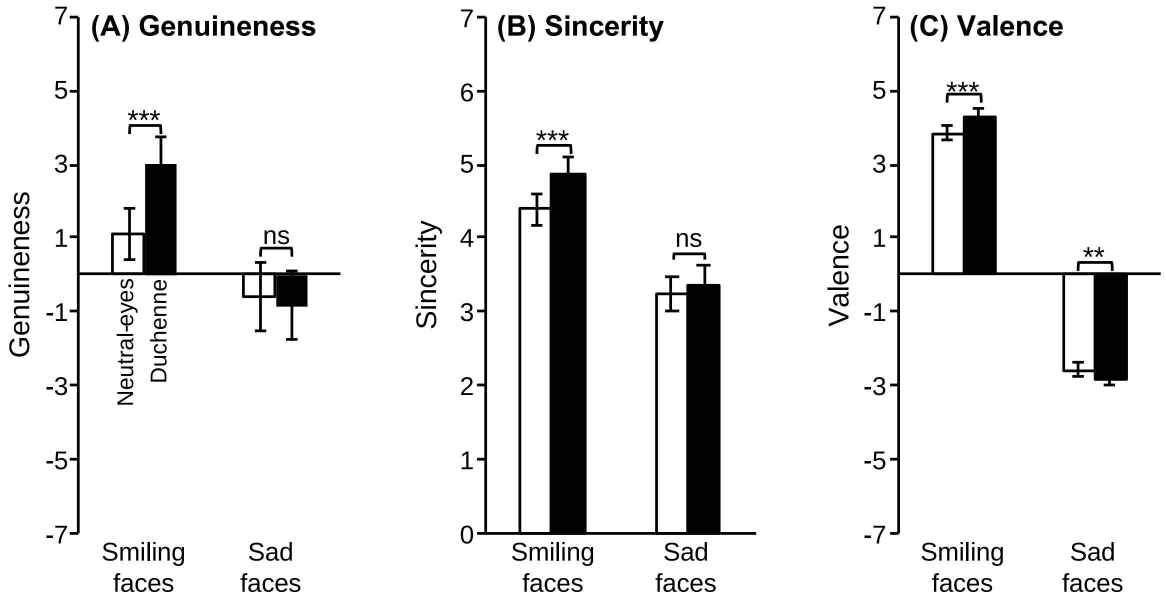

Figure 6. Experiment 3: mean ratings of (A) genuineness, (B) sincerity, and (C) valence for smiling and sad faces. Mean ratings were derived by averaging the participant ratings for each stimulus first, and then averaging across all stimuli. Brackets above each pair of bars indicate results from paired samples *t*-tests. Error bars show 95% CIs for the mean difference from *t*-tests. *p <.05. **p <.01. ***p <.001. ns=non-significant.

Discussion

Experiment 3 replicated our finding in Experiment 2 that the Duchenne marker does not make sad expressions look more genuine relative to neutral-eyes sadness. Here, we found this effect held when tested with Malek et al.'s (2019) sincerity rating task, as well as with our own genuineness rating task, attesting to its robustness. This finding also held in the context of replicating the standard effects found for the Duchenne marker in smiles, which indicates our participants were otherwise responding as expected during the experiment. Overall, our results provide further evidence that the Duchenne marker is not interpreted as a sign of emotional genuineness for sad expressions.

However, we did find Duchenne sad expressions were rated as more negative than neutral-eyes ones. This finding is consistent with Malek et al.'s (2019) results for valence, and suggests the Duchenne marker may have some role in communicating intensity for sad expressions.

Experiment 4: All rating types for Duchenne, neutral-eyes and original sad expressions

Our Experiment 3 finding that Duchenne sad expressions were rated as more negative seems to contradict our earlier Experiment 1 finding they were *not* rated as more sad or physically intense than the comparison sad expressions. However, a critical difference between these two experiments is the comparison expressions. Experiment 1 used original sad expressions, which often included subtle signs of the Duchenne marker and/or AU7. Experiment 3 used neutral-eyes sad expressions, which clearly did not include the Duchenne marker or AU7. Therefore, to elucidate the role of the Duchenne marker in communicating intensity for sad expressions, our final experiment used all of the rating tasks from our earlier experiments (ratings of genuineness, sincerity, sadness, negative valence, and physical intensity) to test perception of Duchenne sad expressions in comparison to both neutral-eyes and original sad expressions. The main comparison of interest here is that between Duchenne and neutral-eyes sad expressions—that is, between expressions that are known to include the Duchenne marker and expressions that are known not to include the Duchenne marker or the closely-related AU7. Original sad expressions were included only to help us understand the apparent difference in our Experiment 1 and 3 findings for intensity.

Method

Participants

Data analysed were from 18 undergraduate students (56% female, 44% male; *M* age=18.72 years, *SD*=0.9, range 18-21) who had not participated in the previous experiments. Participants received course credit for the single 1 hour test session. All other participant details were as for Experiment 1.

Face stimuli

Stimuli were the Duchenne, original, and neutral-eyes sad expressions used in Experiments 1 and 2 (i.e., all 40 KDEF identities).

Experimental procedure (in task order)

Valence ratings. Same as in Experiment 3.

Emotional intensity ratings. Same as in Experiment 1, except participants were only asked to rate the intensity of sadness (i.e., not all six emotions, because Exp. 4 tested only sad expressions).

Sincerity ratings. Same as in Experiment 3.

Genuineness ratings. Same as in Experiment 1, except participants were instructed to rate the genuineness of the *sadness* being shown (i.e., rather than the genuineness of the *emotion* being shown).

Physical intensity ratings. Same as in Experiment 1.

Data analysis

Because participants were asked to rate the genuineness specifically of the sadness being shown, ratings necessarily reflected the perceived genuineness of the intended emotion (i.e., sadness), so there was no need to remove any trials from Experiment 4. Paired-sample *t*-tests were used to compare mean ratings (e.g., of genuineness) for each combination of the three types of sad expressions (i.e., Duchenne vs original sadness, Duchenne vs neutral-eyes sadness, original vs neutral-eyes sadness). A Bonferroni corrected alpha level of .05/3=.017 was applied to adjust for the three comparisons.

Results

For perceptions of genuineness and sincerity, we again found no significant differences between the Duchenne sad and comparison expressions, irrespective of whether the comparison expressions were neutral-eyes or original sad expressions (genuineness ratings: $M_{Duchenne=}$-0.18, $M_{neutral\text{-}eyes=}$-0.21, $t(39)=0.12$, $p=.902$, $d=.02$; $M_{Duchenne=}$-0.18, $M_{original}$ =-0.22, $t(39)=0.18$, $p=.860$, $d=.03$; sincerity ratings: $M_{Duchenne=}$3.31, $M_{neutral\text{-}eyes=}$3.29, $t(39)=0.25$, $p=.807$, $d=.04$; $M_{Duchenne=}$3.31, $M_{original=}$3.35, $t(39)=0.36$, $p=.721$, $d=.06$.

For perceptions of sadness and valence however, we found clear effects when Duchenne sad expressions were compared with neutral-eyes ones, but not when they were compared with original sad expressions. In comparison to neutral-eyes sad expressions, Duchenne sad expressions were rated as significantly more negative, $M_{Duchenne}$=-2.78, $M_{neutral\text{-}eyes}$=-2.35, t(39)=6.45, p<.001, d=1.02, sad, $M_{Duchenne}$=4.94, $M_{neutral\text{-}eyes}$=4.49, t(39)=5.18, p<.001, d=.82, and also physically intense, $M_{Duchenne}$=4.76, $M_{neutral\text{-}eyes}$=4.16, t(39)=7.58, p<.001, d=1.20. In contrast, we found no significant differences between Duchenne and original sad expressions for ratings of valence (how negative), $M_{Duchenne}$=-2.78, $M_{original}$=-2.69, t(39)=1.35, p=.186, d=.21, or sadness, $M_{Duchenne}$=4.94, $M_{original}$=4.85, t(39)=1.23, p=.228, d=.19. Interestingly though, we did find Duchenne sad expressions were rated as more physically intense than original ones this time, $M_{Duchenne}$=4.76, $M_{original}$=4.44, t(39)=3.70, p=.001, d=.58.

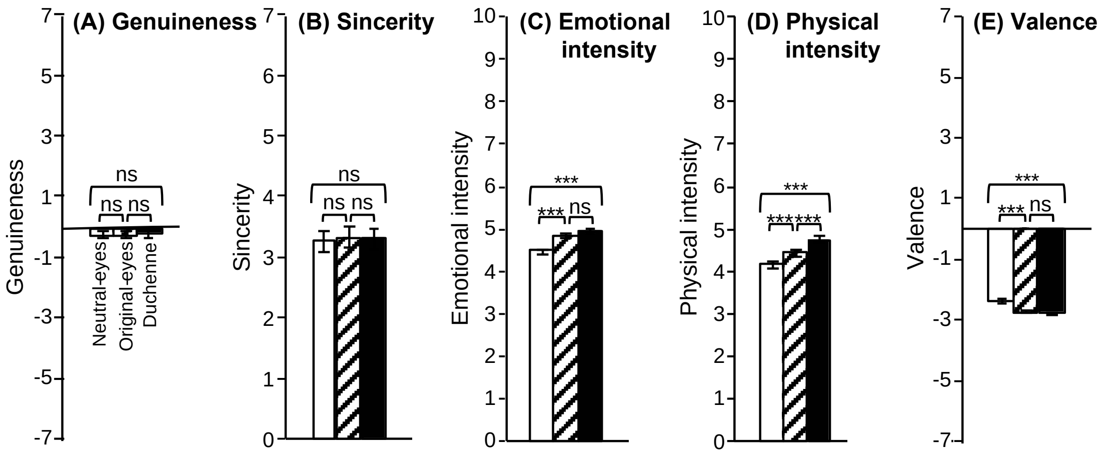

Figure 7. Experiment 4: mean ratings of (A) genuineness, (B) sincerity, (C) emotional (sadness) intensity, (D) physical intensity, and (E) valence for sad faces. Mean ratings were derived by averaging the participant ratings for each stimulus first, and then averaging across all stimuli. Brackets above each pair of bars indicate results from paired samples *t*-tests. Error bars show 95% CIs from repeated-measures ANOVA comparing all three eye-type conditions, calculated using the formula: t-crit*sqrt(MSE/N) (Loftus & Masson, 1994). Note, these ANOVAs produced results that were consistent with those of the *t*-tests. Namely, no significant main effect of condition for genuineness nor sincerity ratings, both *p*s>.87; and significant main effects of condition for emotional intensity, physical intensity, and valence ratings, all *p*s<.001. ****p* <.001. ns=non-significant.

Additional analyses combining data from all four experiments

Given that all four experiments had failed to reveal any effect of the Duchenne marker on the perceived genuineness of sad expressions, we ran two additional analyses to test whether they were able to reveal such an effect. Neither did. First, we looked at results for individual participants. Even with a relaxed criterion of p <.05 *one*-tailed in the predicted direction (i.e., equivalent to p <.10 two-tailed), very few participants rated the Duchenne sad expressions as significantly more genuine than the original or neutral-eyes ones (<5% of participants, compared to >50% of participants for smiling expressions). Second, following the recommendations of Goh, Hall, and Rosenthal (2016), we tested whether cumulating the evidence across our experiments via a mini meta-analysis was able to reveal an effect. Meta-analysis was performed by: (1) calculating Cohen's *dz* using genuineness ratings for each experiment comparing Duchenne sad expressions with: original sadness (Exps. 1, 2 & 4) and, separately, neutral-eyes sadness (Exps. 2, 3 & 4); (2) transforming *dz* values into *r*s; and (3) running two fixed effects models (Model 1=Duchenne vs original-eyes; Model 2=Duchenne vs neutral-eyes) on the *r* values using Fields and Gillets' (2010) Meta_Basic_r.sps syntax in SPSS. Results confirmed our findings from the individual experiments that there was no significant effect of Duchenne status on perceived genuineness for sad expressions: Duchenne vs original sadness, mean r=.066, z=.696, p=.486; and Duchenne vs neutral-eyes sadness, mean r=.011, z=.121, p=.904.

Post-hoc FACS coding by human coders

The presence of the Duchenne marker in the 40 Duchenne smile images was originally verified by automated coding (Calvo, Fernández-Martín, Recio, & Lundqvist, 2018). Following our experiments, we also had all of our 200 stimulus images coded by a human FACS certified coder (Supplement S7). The coder was blind to the experimental results and individually coded the presence and intensity of AU6 and AU7 each image. Intercoder reliability was checked by a second FACS certified coder for 25% of the stimulus material (50 images). Mean agreement for the presence of AU 6 (ICC = .840, Cohen's Kappa (κ) = .837) and AU 7 (ICC = .843, κ = .840) was high. There was also substantial agreement when AU 6 (ICC = .945, κ = .861) and AU7 (ICC = .826, κ = .615) were scored on a 3-point intensity level (x,y,z) in FACS.

This coding revealed three important things. First, as expected, it confirmed that AU6 was *not* present in any of the neutral-eyes expressions. Second, only 31 of the Duchenne smiles and 32 of the Duchenne sad expressions were coded as including the Duchenne marker. We therefore re-ran analyses that compared the human-FACS-confirmed Duchenne smiles (N=31) and sad expressions (*N*=32) with neutral-eyes versions of those expressions. Results fully replicated our findings for the full set of 40 stimulus identities (Supplement S8). Finally, although most of the original sad expressions were found to include AU7 (27 out of 40), only a small number included the Duchenne marker AU6 (only 8 of 40), thereby verifying that the original sad expressions were mostly non-Duchenne sad expressions.

Discussion

Taking the results of Experiment 4 and our additional analysis together with those of our earlier experiments, we now have strong evidence the Duchenne marker does *not* affect the perceived genuineness of sad expressions. Experiment 4 did however find the Duchenne marker caused sad expressions to be perceived as more intense (i.e., sadder, more negative, and more physically intense) than comparison sad expressions that are known not to include the Duchenne marker (AU6) or the closely related AU7 (i.e., both actions clearly not present in neutral-eyes sad expressions). This finding is consistent with previous literature showing the Duchenne marker is associated with greater perceived intensity, for computer-generated sad faces (Malek et al., 2019) and infant cry-faces (Dinehart et al., 2005; Messinger, 2002; Messinger et al., 2012). On the other hand, our Duchenne sad expressions were not perceived as sadder or more negative than original sad expressions. This result supports the concerns we raised following Experiment 1 about the non-neutrality of the eyes in the original sad expressions making them a poor comparison for our Duchenne expressions.

General Discussion

Our results make two main novel empirical contributions to the literature. First, for smiling expressions, we provide the first clear evidence that the association between Duchenne status and perceived genuineness is not fully attributable to intensity. Our finding that the Duchenne marker makes a unique direct contribution to the perceived genuineness of smiles argues against the idea (Fridlund, 1994; Messinger, 2002; Messinger et al., 2012) it should be re-conceptualised as a marker of emotional intensity. Instead, this finding is consistent with the popular notion (e.g., Ekman et al., 1990) that the Duchenne marker signals genuine happiness in smiles, and supports the large body of empirical literature on this topic (e.g., Gunnery & Ruben, 2016; Quadflieg et al., 2013; Song et al., 2016; Wang et al., 2017).

Second, for sad expressions, we found strong evidence across four experiments that the Duchenne marker is *not* interpreted as signalling emotion genuineness. This finding was robust: it held when Duchenne sad expressions were compared with neutral-eyes and original sad expressions; for ratings of genuineness and sincerity, with detailed and minimal instructions respectively; and when data were analysed in a cumulative manner using meta-analysis. Importantly, this lack-of-effect for genuineness occurred in the context of the Duchenne marker having the expected, and previously evidenced (Dinehart et al., 2005; Malek et al., 2019; Messinger, 2002; Messinger et al., 2012), effect on the perceived intensity of sad expressions. Specifically, the Duchenne marker made sad expressions look more intensely sad and negative relative to sad expressions that did not include the Duchenne marker.

Overall, our study provides compelling evidence that perceiver interpretations of the Duchenne marker are not the same for smiles and sad expressions. The Duchenne marker was interpreted as signalling genuineness as well as intensity for smiles, but only intensity for sadness. The common effect of Duchenne status on the perceived intensity of both expressions is consistent with previous work (e.g., Malek et al., 2019; Messinger et al., 2012, Quadfleig et al., 2012) and is perhaps unsurprising. Our Duchenne expressions necessarily included greater facial activation than the neutral-eyes ones, which should logically cause them to be perceived as more intense, at least physically. Importantly though, by separating out emotional from physical intensity in our rating tasks, we were able to establish that Duchenne smiles and sad expressions were perceived as more happy and sad respectively than neutral-eyed versions, not just more physically intense. Indeed, for happiness, the effect size for emotional intensity was notably larger than for physical intensity (e.g., Exp. 1 Duchenne vs neutral-eyes smiles: happiness d=.82, physical intensity d=.40; although this was not the case for sadness: Exp. 4

Duchene vs neutral-eyes sad expressions: sadness d=.82, physical intensity d=1.20). This gain in emotional intensity argues the Duchenne marker has some key role in communicating happiness and sadness. It is already well-established that the Duchenne marker is a core component of happy expressions. For example, Ekman and colleagues consistently identify the critical AUs for happiness as AU6 (the Duchenne marker) and AU12 (smiling) (e.g., Ekman and Friesen, 1982b; Ekman, 2007). In contrast, the Duchenne marker (AU6) is missing from earlier, influential lists of AUs for sad expressions. Most notably, AU6 is missing from the list of AUs for sad expressions in the version of FACS used to code emotional expressions (i.e., in EM-FACS, the AUs for sadness are listed as AU1+4+15; Ekman and Friesen, 1982b). However, more recent lists of AUs for sadness tend to include AU6 (e.g., Ekman, 2007; Keltner, Sauter, Tracy, & Cowen, 2019). The inclusion of the Duchenne marker (AU6) in the list of AUs for sadness is supported by our finding that it makes sad expressions appear sadder, as well as by the evidence reviewed in our Introduction showing the Duchenne marker regularly appears in expressions of sadness (e.g., Keltner & Bonanno, 1997; Lee & Beattie, 1998; Namba et al., 2017).

An interesting question raised by our findings is why the Duchenne marker influences the perceived genuineness of smiles but not sad expressions. While our current data cannot answer this question, we suspect people may only learn to associate Duchenne smiling with genuine happiness through exposure. That is, by repeatedly seeing people display intense smiles, which are prone to include the Duchenne marker (Ekman & Friesen, 1982a; Krumhuber & Manstead, 2009) in contexts where they expect the displayer to be feeling genuinely happy. This learning hypothesis is suggested by evidence that people from some cultures do not interpret the Duchenne marker as a sign of genuine happiness within their own cultural group, but learn to interpret it as a sign of genuine happiness in people from other cultures when living abroad (Thibault et al., 2012). Also, in cultures that do interpret the Duchenne marker as a sign of genuine happiness, this effect does not emerge until around 14 years of age (in Australia, Dawel, Palermo, O'Kearney, & McKone, 2015; and in Canada, Thibault, Gosselin, Brunel, & Hess, 2009).

A second issue that remains unresolved is why we failed to replicate Malek et al.'s (2019) finding that Duchenne sadness was perceived as more sincere than neutral-eyes sadness. One important methodological difference is that we used 40 human face identities and manipulated their expressions digitally, whereas Malek et al. (2019) used three FaceGen identities and manipulated their expressions using FACSGen software (Krumhuber et al., 2012). It is possible our digital transplants looked somehow artificial. However, no participant reported noticing this manipulation when asked about the stimuli. Also, if the transplants were the problem, the Duchenne sad expressions should have looked *less* genuine and sincere than the original ones, and this was not the case. Certainly, for smiles, manipulating images digitally like we did has produced findings for genuineness that are similar to those from naturally-occurring expressions (Gunnery & Ruben, 2016; Quadflieg et al., 2013; Thibault et al., 2009). An alternative explanation is that participants rated Malek et al.'s (2019) Duchenne sad expressions as more sincere because they perceived these more expressive computer-generated images as being more likely to have a mind (i.e., increased animacy) than the ones with neutral eyes. This argument draws on recent evidence that facial expressivity can increase the perceived animacy of face images, including for computer-generated faces (Krumhuber, Lai, Rosin, & Hugenberg, 2019).

A strength of our digital transplantation procedure is that we can be sure the only differences between the faces in different conditions are the ones around the eyes that we are interested in (i.e., we know all of the other muscles in the face appear activated to the same degree). However, it is possible the Duchenne marker might appear differently in naturalistic expressions of sadness than it did in our stimuli. Although the Duchenne marker often appears in combination with the closely related AU7 in happy as well as sad expressions (Del Giudice & Colle, 2007; Scherer et al., 2019), there could be subtle differences in the activation pattern between the two emotions. For example, activation of the Duchenne marker might differentially affect how the eyebrows appear in smiles and sadness. For smiles, there is no specific involvement of the eyebrows. In contrast, sadness is expressed in part by raising the inner part of the eyebrows (AU1) in combination with lowering the outer part (AU4) (Ekman, 2007). Duchenne (AU6) activation lowers the outer eyebrows in a way that is consistent with this AU1+4 combination, but still the physical Duchenne information we extracted from happy expressions might have appeared somewhat differently to how it does in sadness. Thus it is possible the way in which the Duchenne marker appears in naturally-occurring sad expressions might influence perceived genuineness. Ideally, future work should compare naturally-occurring sad expressions that do and do not include the Duchenne marker, but are matched for other relevant features, including identity and level of physical activation of other AUs to the degree that this is possible. Finally, we emphasise our findings relate only to the *perception* of facial expressions. It is possible studies that examine how people feel when naturally producing facial expressions of sadness might uncover some link between feeling genuinely sad and the Duchenne marker. It will be important for future work to examine the Duchenne marker in the production as well as perception of naturally occurring sad expressions.

Acknowledgements

We thank Xia Fang for FACS coding, and Rachel Robbins and the three anonymous reviewers for their feedback on earlier drafts of this manuscript.

References

Aviezer, H., Zangvil, S., Messinger, D. S., Mattson, W. I., Gangi, D. N., & Todorov, A. (2015). Thrill of victory or agony of defeat? Perceivers fail to utilize information in facial movements. *Emotion, 15*(6), 791-797. doi: 10.1037/emo0000073

Calvo, M. G., & Fernández-Martín, A. (2013). Can the eyes reveal a person's emotions? Biasing role of the mouth expression. *Motivation and Emotion, 37*(1), 202-211. doi: 10.1007/s11031-012-9298-1

Calvo, M. G., Fernández-Martín, A., Recio, G., & Lundqvist, D. (2018). Human observers and automated assessment of dynamic emotional facial expressions: KDEF-dyn database validation. *Frontiers in psychology, 9*, 2052. doi: 10.3389/fpsyg.2018.02052

Carroll, J. M., & Russell, J. A. (1997). Facial Expressions in Hollywood's Portrayal of Emotion. *Journal of Personality and Social Psychology, 72*(I), 164-176. doi: 10.1037//0022-3514.72.1.164

Darwin, C. (1872/2009). The expression of emotion in man and animals In F. Darwin (Ed.), *The expression of the emotions in man and animals, anniversary edition* (pp. 33-307). London: John Murray.

Dawel, A., Palermo, R., O'Kearney, R., & McKone, E. (2015). Children can discriminate the authenticity of happy but not sad or fearful facial expressions, and use an immature intensity-only strategy. *Frontiers in Psychology, 6*(462), 1-16. doi: 10.3389/fpsyg.2015.00462

Dawel, A., Wright, L., Irons, J., Dumbleton, R., Palermo, R., O'Kearney, R., & McKone, E. (2017). Perceived emotion genuineness: normative ratings for popular facial expression stimuli and the development of perceived-as-genuine and perceived-as-fake sets. *Behavior Research Methods, 49*(4), 1539-1562. doi: 10.3758/s13428-016-0813-2

DeBruine, L. & Jones, B. (2018). Determining the number of raters for reliable mean ratings. Retrieved from https://debruine.github.io/posts/how-many-raters/

Del Giudice, M., & Colle, L. (2007). Differences between children and adults in the recognition of enjoyment smiles. *Developmental Psychology, 43*(3), 796-803. doi: 10.1037/0012-1649.43.3.796

Dinehart, L., Messinger, D., Acosta, S., Cassel, T., Ambadar, Z., & Cohn, J. (2005). Adult perceptions of positive and negative infant emotional expressions. . *Infancy, 8*(3), 279-303. doi: 10.1207/s15327078in0803_5

Duchenne de Boulogne, G. B. (1862/1990). *The Mechanism of Human Facial Expression*. New York: Cambridge University Press.

Ekman, P. (2007). The directed facial action task: Emotional responses without appraisal. In Coan, J.A. & Allen, J.J.B. (Eds), *Handbook of emotional elicitation and assessment* (p. 47-53). New York: Oxford University Press.

Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The Duchenne smile: Emotional expression and brain physiology: II. *Journal of Personality and Social Psychology, 58*(2), 342-353. doi: 10.1037/0022-3514.58.2.342

Ekman, P., & Friesen, W. V. (1982a). Felt, false, and miserable smiles. *Journal of Nonverbal Behavior, 6*(4), 238-252. doi: 10.1007/BF00987191

Ekman, P., & Friesen, W. V. (1982b). EMFACS. Unpublished manuscript, San Francisco, CA.

Ekman, P., & Friesen, W. V. (1988). Smiles when lying. *Journal of Personality and Social Psycology, 54*(3), 414-420. doi: 10.1037/0022-3514.54.3.414

Ekman, P., Friesen, W. V., & Hagar, J. C. (2002). *Facial Action Coding System: The Manual on CD Rom* Salt Lake City, UT: Research Nexus Division of Network Information Research Corporation.

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*(4), 1149-1160. doi: 10.3758/Brm.41.4.1149

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191. doi: Doi 10.3758/Bf03193146

Frank, M. G., Ekman, P., & Friesen, W. V. (1993). Behavioral Markers and Recognizability of the Smile of Enjoyment. *Journal of Personality and Social Psychology, 64*(1), 83-93. doi: 10.1037/0022-3514.64.1.83

Fridlund, A. (1994). *Human Facial Expression: An Evolutionary View*. San Diego: Academic Press, Inc.

Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini Meta-Analysis of Your Own Studies: Some Arguments on Why and a Primer on How. *Social and Personality Psychology Compass, 10*(10), 535-549. doi: 10.1111/spc3.12267

Grogorick, S., Albuquerque, G., Tauscher, J.-P., Kassubeck, M., & Magnor, M. (2019). *Towards VR Attention Guidance: Environment-dependent Perceptual Threshold for Stereo Inverse Brightness Modulation*. Barcelona, Spain.

Gunnery, S. D., Hall, J. A., & Ruben, M. A. (2013). The Deliberate Duchenne Smile: Individual Differences in Expressive Control. *Journal of Nonverbal Behavior, 37*(1), 29-41. doi: 10.1007/s10919-012-0139-4

Gunnery, S. D., & Ruben, M. A. (2016). Perceptions of Duchenne and non-Duchenne smiles: A meta-analysis. *Cognition and Emotion, 30*(3), 501-515. doi: 10.1080/02699931.2015.1018817

Harris, C. R., & Alvarado, N. (2005). Facial expressions, smile types, and self-report during humour, tickle, and pain. *Cognition and Emotion, 19*(5), 655-669. doi: 10.1080/02699930441000472

Hayes, A. F. (2018). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach* (Second ed.). New York: Guilford Press.

Hess, U., & Bourgeois, P. (2010). You smile-I smile: Emotion expression in social interaction. *Biological Psychology, 84*(3), 514-520. doi: 10.1016/j.biopsycho.2009.11.001

Keltner, D., & Bonanno, G. (1997). A Study of Laughter and Dissociation: Distinct Correlates of Laughter and Smiling During Bereavement. *Journal of Personality and Social Psychology, 73*(4), 687-702.

Keltner, D., Sauter, D., Tracy, J., & Cowen, A. (2019). Emotional expression: Advances in basic emotion theory. *Journal of Nonverbal Behavior*, *43*(2), 133-160. doi:10.1007/s10919-019-00293-3

Khan, M. M., Ward, R. D., & Ingleby, M. (2009). Capturing physiology of emotion along facial muscles: A method of distinguishing feigned from involuntary expressions. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 5702 LNCS*, 1196-1203. doi: 10.1007/978-3-642-03767-2_145

Knoll, T., & Knoll, J. (2018). Adobe Photoshop CC (Version 2017.1.0). Retrieved from https://www.adobe.com/au/products/photoshop.html

Krumhuber, E. G., Lai, Y. K., Rosin, P. L., & Hugenberg, K. (2019). When Facial Expressions Do and Do Not Signal Minds: The Role of Face Inversion, Expression Dynamism, and Emotion Type. *Emotion*, *19*(4), 746-750. doi: 10.1037/emo0000475

Krumhuber, E. G., & Manstead, A. S. R. (2009). Can Duchenne Smiles Be Feigned? New Evidence on Felt and False Smiles. *Emotion, 9*(6), 807-820. doi: 10.1037/a0017844

Krumhuber, E. G., Tamarit, L., Roesch, E. B., & Scherer, K. R. (2012). FACSGen 2.0 Animation Software: Generating Three-Dimensional FACS-Valid Facial Expressions for Emotion Research. *Emotion*, *12*(2), 351-363. doi: 10.1037/a0026632

Lee, V., & Beattie, G. (1998). The rhetorical organization of verbal and nonverbal behavior in emotion talk *Semiotica*, *120*(1), 39-92. doi: https://doi.org/10.1515/semi.1998.120.1-2.39

Lundqvist, D., Flykt, A., & Öhman, A. (1998a). *The Karolinska directed emotional faces - KDEF (CD ROM)*. Stockholm: Karolinska Institutet, Department of Clinical Neuroscience, Psychology section.

Lundqvist, D., Flykt, A., & Öhman, A. (1998b). About KDEF, from http://kdef.se/home/aboutKDEF.html

Malek, N., Messinger, D., Gao, A. Y. L., Krumhuber, E., Mattson, W., Joober, R., . . . Martinez-Trujillo, J. C. (2019). Generalizing Duchenne to Sad Expressions with Binocular Rivalry and Perception Ratings. *Emotion, 19*(2), 234-241. doi: 10.1037/emo0000410

Matsumoto, D., & Willingham, B. (2009). Spontaneous Facial Expressions of Emotion of Congenitally and Noncongenitally Blind Individuals. *Journal of Personality and Social Psychology, 96*(1), 1-10. doi: 10.1037/a0014037

Mattson, W., Cohn, J. F., Mahoor, M. H., Gangi, D. N., & Messinger, D. (2013). Darwin's Duchenne: Eye Constriction during Infant Joy and Distress. *PLoS ONE, 8*(11), 1-5. doi: 10.1371/journal.pone.0080161

Mehu, M., Little, A. C., & Dunbar, R. I. M. (2007). Duchenne smiles and the perception of generosity and sociability in faces. *Journal of Evolutionary Psychology, 5*(1), 183-196. doi: 10.1556/JEP.2007.1011

Mehu, M., Mortillaro, M., Bänziger, T., & Scherer, K. R. (2012). Reliable facial muscle activation enhances recognizability and credibility of emotional expression. *Emotion, 12*(4), 701-715. doi: 10.1037/a0026717

Messinger, D. S. (2002). Positive and negative: Infant facial expressions. *Current Directions in Psychological Science, 11*(1), 1-6. doi: 10.1111/1467-8721.00156

Messinger, D. S., Mattson, W. I., Mahoor, M. H., & Cohn, J. F. (2012). The eyes have it: Making positive expressions more positive and negative expressions more negative. *Emotion, 12*(3), 430-436. doi: 10.1037/a0026498

Namba, S., Kagamihara, T., Miyatani, M., & Nakao, T. (2017). Spontaneous Facial Expressions Reveal New Action Units for the Sad Experiences. *Journal of Nonverbal Behavior, 41*(3), 203-220. doi: 10.1007/s10919-017-0251-6

Papa, A., & Bonanno, G. (2008). Smiling in the Face of Adversity: The Interpersonal and Intrapersonal Functions of Smiling. *Emotion, 8*(1), 1-12. doi: 10.1037/1528-3542.8.1.1

Peirce, J. W. (2007). PsychoPy--Psychophysics software in Python. *J Neurosci Methods, 162*(1-2), 8-13. doi: 10.1016/j.jneumeth.2006.11.017

Peirce, J. W. (2008). Generating Stimuli for Neuroscience Using PsychoPy. *Front Neuroinform, 2*, 10. doi: 10.3389/neuro.11.010.2008

Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, *104*(1), 1-15. doi: Doi 10.1016/S0001-6918(99)00050-5

Quadflieg, S., Vermeulen, N., & Rossion, B. (2013). Differential Reliance on the Duchenne Marker During Smile Evaluations and Person Judgments. *Journal of Nonverbal Behavior*, *37*(2), 69-77. doi: 10.1007/s10919-013-0147-z

Scherer, K. R., & Ellgring, H. (2007). Are facial expressions of emotion produced by categorical affect programs or dynamically driven by appraisal? *Emotion, 7*(1), 113-130. doi: 10.1037/1528-3542.7.1.113

Scherer, K. R., Ellgring, H., Dieckmann, A., Unfried, M., & Mortillaro, M. (2019). Dynamic Facial Expression of Emotion and Observer Inference. *Frontiers in Psychology, 10*. doi: ARTN 50810.3389/fpsyg.2019.00508

Schneider, K., & Josephs, I. (1991). The expressive and communicative functions of preschool childrens's smiles in an achievement-situation. *Journal of Nonverbal Behavior, 15*(3), 185-198.

Schneider, K., & Unzer, L. (1992). Preschoolers' attention and emotion in an achievement and an effect game: A longitudinal study. *Cognition & Emotion, 6*(1), 37-63. doi: 10.1080/02699939208411057

Song, R., Over, H., & Carpenter, M. (2016). Young children discriminate genuine from fake smiles and expect people displaying genuine smiles to be more prosocial. *Evolution and Human Behavior, 37*(6), 490-501. doi: 10.1016/j.evolhumbehav.2016.05.002

Sorokowska, A., Sorokowski, P., Hilpert, P., Cantarero, K., Frackowiak, T., Ahmadi, K., . . . Pierce, J. D. (2017). Preferred Interpersonal Distances: A Global Comparison. *Journal of Cross-Cultural Psychology, 48*(4), 577-592. doi: 10.1177/0022022117698039

Thibault, P., Gosselin, P., Brunel, M.-l., & Hess, U. (2009). Children's and adolescents' perception of the authenticity of smiles. *Journal of Experimental Child Psychology, 102*(3), 360-367. doi: 10.1016/j.jecp.2008.08.005

Thibault, P., Levesque, M., Gosselin, P., & Hess, U. (2012). The Duchenne Marker is Not a Universal Signal of Smile Authenticity – But it Can Be Learned ! *Social Psychology, 43*(4), 215-221. doi: 10.1027/1864-9335/a000122

Wang, Q., Xu, Z., Cui, X., Wang, L., & Ouyang, C. (2017). Does a big Duchenne smile really matter on e-commerce websites? An eye-tracking study in China. *Electronic Commerce Research, 17*(4), 609-626. doi: 10.1007/s10660-016-9237-4

Weinberg, S. M., Raffensperger, Z. D., Kesterke, M. J., Heike, C. L., Cunningham, M. L., Hecht, J. T., . . . Marazita, M. L. (2016). The 3D facial norms database: Part 1. A web-based craniofacial anthropometric and image repository for the clinical and research community. *The Cleft Palate-Craniofacial Journal, 53*(6), 185-197. doi:10.1597/15-199