Journal of Deaf Studies and Deaf Education, 2019, 1–10

doi: 10.1093/deafed/enz023
Empirical Manuscript

EMPIRICAL MANUSCRIPT

Quantifying Facial Expression Intensity and Signal Use in Deaf Signers

Chloé Stoll[1,2,‡], Helen Rodger[3,†,‡], Junpeng Lao[3,‡], Anne-Raphaëlle Richoz[3], Olivier Pascalis[1], Matthew Dye[4], and Roberto Caldara[3,*]

[1]Laboratoire de Psychologie et de Neurocognition (CNRS-UMR5105), Université Grenoble-Alpes, [2]Laboratory for Investigative Neurophysiology, Centre Hospitalier Universitaire Vaudois and University of Lausanne, [3]Eye and Brain Mapping Laboratory (iBMLab), Department of Psychology, University of Fribourg and [4]National Technical Institute for Deaf/Rochester Institute of Technology

*Correspondence should be sent to: Roberto Caldara, Department of Psychology, University of Fribourg, R. Faucigny 2, 1700 Fribourg, Switzerland (e-mail: roberto.caldara@unifr.ch)

Abstract

We live in a world of rich dynamic multisensory signals. Hearing individuals rapidly and effectively integrate multimodal signals to decode biologically relevant facial expressions of emotion. Yet, it remains unclear how facial expressions are decoded by deaf adults in the absence of an auditory sensory channel. We thus compared early and profoundly deaf signers ($n = 46$) with hearing nonsigners ($n = 48$) on a psychophysical task designed to quantify their recognition performance for the six basic facial expressions of emotion. Using neutral-to-expression image morphs and noise-to-full signal images, we quantified the *intensity* and *signal* levels required by observers to achieve expression recognition. Using Bayesian modeling, we found that deaf observers require more signal and intensity to recognize disgust, while reaching comparable performance for the remaining expressions. Our results provide a robust benchmark for the intensity and signal use in deafness and novel insights into the differential coding of facial expressions of emotion between hearing and deaf individuals.

As a social species, one crucial ability for survival is the effective communication and decoding of social information. We acquire the ability to understand others and express our inner feelings long before language develops. Indeed, nonverbal communication is a major part of our social, interpersonal interaction. It conveys a rich set of information, which is at times beyond the limits of human language. One such social information code is revealed by our facial expressions (Jack & Schyns, 2015), which are influenced by culture (Jack, Blais, Scheepers, Schyns, & Caldara, 2009; Jack, Caldara, & Schyns, 2012a; Jack, Garrod, Yu, Caldara, & Schyns, 2012b—for a review see Caldara, 2017) from an early stage of life (Geangu et al., 2016). The ability to decode emotional cues in our social environment is essential for normative social functioning. Effective representations of our own and others' internal affective states modeled through facial expressions, and accurately expressing these states through facial muscle movements, plays a central role in defining healthy social relations and well-being (e.g. Carton, Kessler, & Pape, 1999; Feldman, Philippot, & Custrini, 1991; Izard et al., 2001; Nowicki & Duke, 1992).

The ability to decode and recognize facial expressions of emotion has been little studied in deaf individuals, considered by many as a distinct cultural group (Jones, 2002). For deaf people, faces have a special status when it comes to communication.

†Helen Rodger, http://orcid.org/0000-0002-8463-9551
‡These first authors contributed equally to this work.

Received October 9, 2018; revisions received April 30, 2019; accepted May 3, 2019

For example, in sign language, facial expression conveys not only emotional but also grammatical and syntactic information (Brentari & Crossley, 2002; Liddell, 2003; Reilly, Mcintire, & Seago, 1992; Reilly & Bellugi, 1996). The face and its expressions can also function as intensity markers, and, as a result, the same sign can have different meanings depending on the accompanying facial expression. Sign language communication therefore requires a specific ability to differentiate between syntactic and emotional facial expressions. Moreover, in deaf signers, syntactic and emotional facial expressions are processed by different cortical regions (Corina, Bellugi, & Reilly, 1999; McCullough, Emmorey, & Sereno, 2005). Surprisingly, as might otherwise have been predicted, Grossman and Kegl (2007) did not observe better categorization performance in deaf signers compared to hearing nonsigners for both emotional and linguistic dynamic facial expressions. Instead, they found that hearing nonsigners were better than deaf signers at categorizing facial expressions from interrogative *yes/no* and *wh questions* (e.g. where? what? who? when?) in American Sign Language (ASL). Also, surprisingly, another study (Campbell, Woll, Benson, & Wallace, 1999) has shown that users of British Sign Language (BSL; deaf or hearing) were not better than hearing nonsigners in categorizing facial expressions from BSL interrogative sentences (i.e. yes/no versus wh- question), or more common facial expressions from both BSL and emotional expressions (e.g. puzzled and surprised faces).

The absence of voice tone information may also change a deaf individual's representation of emotional expressions. An emotional response to a specific event often includes an association between a facial expression and a sound. For example, the expression of fear is frequently associated with a loud vocal expression. Fear already captures attention efficiently at a very early age (e.g. Bayet et al., 2017) and elicits stronger identity neural coding in adults compared to other expressions from the visual signal alone (e.g. Turano et al., 2017). In the classic shower scene from Alfred Hitchcock's *Psycho* (1960), a hearing person will immediately associate the terrified face of Janet Leigh with the loud scream heard seconds before she is stabbed in the shower. Different studies have indeed shown that multisensory information is integrated during the processing of facial expressions of emotion (Campanella & Belin, 2007; Collignon et al., 2008) and that this multisensory integration is likely to undergo perceptual narrowing (Lewkowicz, 2014; Lewkowicz & Ghazanfar, 2009).

Therefore, in the absence of auditory information, it is possible that there is a difference in the development of and sensitivity to facial expressions between deaf and hearing individuals. However, in an assessment of the emotional valence of different static facial expressions, Watanabe, Matsuda, Nishioka, and Namatame (2011) did not observe any differences in the judgments made by deaf and hearing participants. In another study using arguably more ecological stimuli (Jones, Gutierrez, & Ludlow, 2017), the authors explored emotion recognition in deaf and hearing children for both static and moving faces (Study 1) and for different intensities of emotion in static images (Study 2). Deaf children recognized facial expressions better in the dynamic compared to the static condition, whereas no difference was found between conditions for the control group of hearing children. Moreover, both groups benefited similarly in terms of better performance when the expressions were higher in emotional intensity. In both Studies 1 and 2, deaf and hearing children showed similar performance as a function of emotion categories, with the exception of disgust, for which deaf children had fewer correct responses. The differences in results between these developmental studies are most likely based on task differences as Watanabe et al. (2011) investigated emotional valence, whereas Jones et al. (2017) investigated emotion recognition. However, similar emotion recognition tests with adult populations have not revealed any differences in recognition between deaf and hearing adult observers (Grossman & Kegl, 2007). The absence of differences in the adult population could be because the mental representations used to decode facial expressions of emotion are similar in both deaf and hearing adults, or the measures used in previous studies could lack sufficient sensitivity to uncover any differences, at least within the adult population.

To obtain a sensitive measure of facial expression recognition performance, we introduced a psychophysical approach. This approach provides a precise measure of recognition performance as the quantity of *signal* (a facial expression of emotion at its fullest intensity modified with random image noise) or *intensity* (a neutral to full intensity emotional expression) is parametrically manipulated. Quantity refers to the signal-to-noise ratio, with high signal quantity meaning a facial expression of emotion with very little added noise, and low signal quantity meaning higher levels of noise. Intensity refers to the amount of emotion coded in the image relative to a neutral baseline. On a continuum from a neutral to a fearful facial expression, images from the fearful end of the continuum would have high emotional intensity. We predicted that, in the absence of auditory information, there will be a difference in the sensitivity to facial expressions between deaf and hearing individuals. Specifically, we predicted that deaf signers would have greater sensitivity to the facial expression visual stimuli used in the signal and intensity tasks and therefore show better overall recognition performance than hearing nonsigners. We tested recognition performance for what are commonly referred to as the six basic emotions. The basic emotions (i.e. anger, fear, disgust, happiness, sadness, and surprise) have been described as basic because they are thought to be universally recognized, although this belief has now been strongly contested (e.g. Crivelli, Russell, Jarillo, & Fernández-Dols, 2016; Gendron, Roberson, van der Vyver, & Barrett, 2014; Jack et al., 2012a; Jack, Sun, Delis, Garrod, & Schyns, 2016). However, these six expressions are the most widely studied in facial expression recognition research and were therefore selected for this study. Due to the mixed results on recognition performance in deaf and hearing cohorts described in the literature earlier, we did not have a prediction that performance between these groups would be better for a specific emotion. We uniquely predicted, as described above, that deaf signers would have greater sensitivity to the facial expression stimuli used in the signal and intensity tasks and therefore show better recognition performance than hearing nonsigners.

To achieve this goal, in the current study we combined a threshold computation paradigm (i.e. QUEST) with a hierarchical mixed-effect model. Using QUEST, we can accurately estimate individual recognition performance for each facial expression of emotion. Moreover, QUEST also returns for each individual the uncertainty of this performance estimation, which is often discarded in data analysis. Here, to properly account for the measurement error from QUEST (i.e. the uncertainty of the performance estimation), we applied a Bayesian Hierarchical mixed-effect model. Although the overall model is nearly identical to a linear mixed model that is commonly used in behavioral research, by adopting a Bayesian approach we can directly account for the noise measurement at the individual level. As a result, we can get more realistic and unbiased group estimations

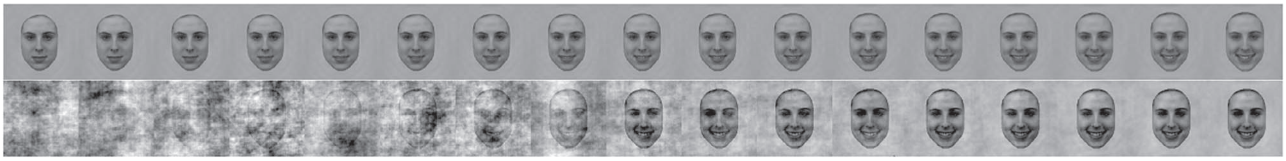

Figure 1 Example stimuli from the signal and intensity conditions. In this image, the stimuli are shown in 5% increments, starting at 20% signal or intensity.

of the latent (true) recognition performance. The full modeling details can be found in the open source analysis code.

Method and Analysis

Participants

We tested a total of 46 deaf signers (26 females), who all had severe to profound hearing loss (dB loss >70) from birth or with onset in the first years of life, and who were all native or early ASL signers (exposed at birth or before the age of 5 years). Among the deaf participants, 12 used cochlear implants (4 occasionally, 8 all the time/every day) and 12 used a hearing aid (7 occasionally, 5 all the time/every day). In addition, 19 participants had a deaf family member (parents and/or siblings). The age range of the deaf participants was between 18 and 30 years (the mean age was 21.70 with a standard deviation of 2.35). In addition, 48 hearing participants (29 females) with no knowledge of any sign language were tested. The age range of the hearing participants was between 18 and 31 years (the mean age was 21.44 with a standard deviation of 3.30). All participants had normal or corrected-to-normal vision. All participants were students attending the Rochester Institute of Technology and the National Technical Institute for the Deaf and received $10 for their participation. The local Institutional Review Board at Rochester Institute of Technology approved this study, and all participants provided written informed consent.

Materials

The stimuli and paradigm were the same as those previously used to test recognition differences in typically developing children (Rodger, Lao, & Caldara, 2018). Facial stimuli expressing each of the six basic emotions (anger, fear, disgust, happiness, sadness, and surprise) and a neutral expression were selected from the Karolinska Directed Emotional Faces database (KDEF, Lundqvist, Flykt, & Öhman, 1998). Images were scaled to 256×256 pixels and mapped to grey scale. For the intensity condition, eight identities (four females) were chosen. We used Abrosoft FantaMorph software to create morphs of 100 increments for each identity and emotional expression, ranging from a 1% morph of a neutral face to a 100% expressive face (original image). The total number of images used was 4,800 (8 identities $\times$ 6 expressions $\times$ 100 increments).

For the signal condition, the stimuli consisted of 252 images from the KDEF database comprising 36 distinct identities (18 females) each displaying 6 facial expressions and a neutral one. Example stimuli of different expression intensities and signal strengths are shown in Figure 1. Participants only viewed images at the intensities or signal strengths calculated by the QUEST (Watson & Pelli, 1983) procedure (see below).

All images were cropped around the face to remove distinctive hairstyles using Adobe Photoshop and were aligned along the eyes and mouth using Psychomorph software (Tiddeman, Burt, & Perrett, 2001). Images were also normalized for contrast and luminance using the SHINE toolbox (Willenbockel et al., 2010) in MATLAB 7.10.0 and displayed on an 800×600 grey background at a distance of 50 cm subtending $10° \times 14°$ of visual angle to simulate a natural viewing distance during social interaction (Hall, 1966). The stimuli were presented using the Psychophysics toolbox (PTB-3) with MATLAB 7.10.0 and QUEST, a Bayesian adaptive psychometric method (described below) to estimate the level of stimuli strength (intensity or signal) for each trial.

Procedure

Participants were asked to decide how the person in each image was feeling by pressing one of six keys on the keyboard each corresponding to one of the emotion categories. We instructed the participants to respond as accurately as they could, as reaction time was not important for the current task. Participants could also press the space bar to indicate "I don't know or I'm uncertain". Participants were given as much time as they needed to familiarize themselves with the response keys before beginning the experiment. Unknown to the participants, the first six trials were practice trials to familiarize the procedure and did not count toward the final threshold estimations. The experimental trials therefore followed seamlessly without distinction from the practice trials. Each trial began with a fixation cross for 500 ms, followed by a face stimulus presented for 500 ms. The displayed intensity or signal strength was provided by the QUEST psychometric procedure (described below) and followed by a mask of random noise until a response was given (see Figure 2 for an illustrated example of a trial). The emotional expressions were randomly displayed in each trial. Once the threshold of a specific expression was estimated by the QUEST procedure, that particular expression was no longer displayed and only the remaining expressions were presented to the participant. As a consequence, the number of trials for each participant varied as a function of the QUEST procedure. The intensity and signal conditions were randomized for each participant, and the mean recognition thresholds were estimated separately. The whole experiment lasted about 30 min. For deaf participants, instructions were both written and signed by the experimenter.

The QUEST Bayesian Adaptive Psychometric Procedure. We implemented the same QUEST procedure as previously used in Rodger, Vizioli, Ouyang, and Caldara (2015) and Rodger et al. (2018). It is an adaptive staircase procedure based on a psychometric function to establish an observer's threshold sensitivity to some physical measure of a stimulus (e.g. stimulus strength, Watson & Pelli, 1983). The obtained threshold can be considered as a measure of how effectively an observer can discriminate a set of stimuli. Here, we tested threshold sensitivity for (1) intensity and (2) signal of emotional expressions between deaf signers and hearing nonsigners. QUEST obtains thresholds by adaptively presenting a sequence of stimuli according to the observer's previous responses. For example, if the observer incorrectly labels the expression of happiness, the subsequent presentation of

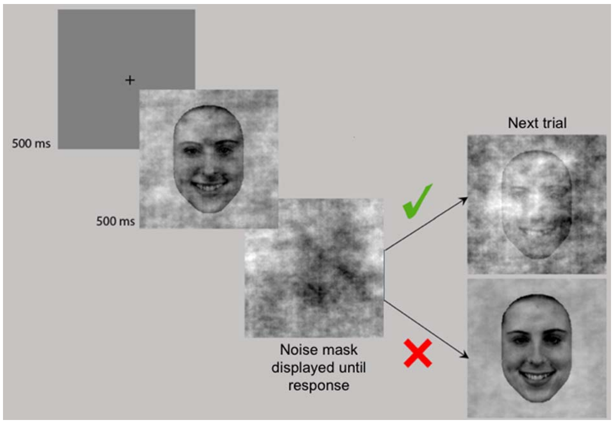

Figure 2 Example trial from the signal condition. Each trial begins with a fixation cross for 500 ms, followed by a face stimulus presentation for 500 ms. The displayed signal or intensity strength is provided by the QUEST psychometric procedure, followed by a mask of random noise until a response is made. Depending on accuracy, the next trial was followed by a face containing more (in the case of an erroneous response) or less (in the case of an accurate response) signal or intensity.

happiness will contain more signal or intensity. Conversely, if the expression is correctly labeled, the next presentation of the same expression will contain less signal or intensity. Thus, this adaptive staircase method is an efficient way to determine the individual perceptual threshold for a type of stimulus (here facial expression), as the represented stimuli are adaptively narrowed to the true underlying threshold.

The QUEST paradigm was implemented in MATLAB 7.10.0 with the Psychophysics Toolbox (PTB-3). We extended the original binary response in QUEST to parametrically determine an observer's perceptual threshold for discriminating each of the six emotional expressions. The final estimated threshold is determined as the intensity or signal strength when the participant maintains performance at 75% accuracy. In this way, equal performance is maintained across observers. For the intensity condition, we implemented one QUEST procedure with an initial expression intensity of 30%. This intensity was selected because by nature, 50% intensity denotes an image morph of 50% neutral and 50% expressive face stimuli, so the initial value should be below this level of morph. For the signal condition, we implemented three QUEST procedures with different initial stimulus strengths (60%, 40%, and 20% signal) to prevent possible bias in the final estimation toward the direction of the initial value. The QUEST procedure terminated for an expression after three consecutive correct or incorrect trials in which the intensity or signal strength standard deviations were less than .025.

Threshold Detection

To calculate the individual thresholds obtained by the QUEST procedure, we computed the mean and the standard deviation of the estimated threshold (Sims & Pelli, 1987; King-Smith, Grigsby, Vingrys, Benes, & Supowit, 1994). In the signal task where multiple QUEST procedures were employed, we computed the arithmetic mean to get the final threshold estimation and the quadratic mean of the standard deviations to get the error of the estimated threshold (based on the Gaussian assumption of the estimation). In a previous paper (Rodger et al., 2015) we used the intensity of the last trial from the QUEST procedure as the threshold estimation. However, because some participants could not achieve 75% recognition performance even when intensity or signal was at the maximum (100%), the previous calculation returned a ceiling value of 100% (especially for some expressions, e.g. fear; see supplementary figure in Rodger et al., 2015). Using the new computation approach, the threshold instead returned values above 100%.

Statistical Modeling

Data analysis was performed in Python using Jupyter Notebook. Bayesian modeling was performed using PyMC3 version 3.6, and the results were displayed using Seaborn and Matplotlib. The full model parameterization is displayed below:

$$
\begin{aligned}
&\text{Intercept}\\
&\qquad C \sim \text{Normal}\,(0.5, 1)\\
&\text{Coefficient (for Fixed effect)}\\
&\qquad \beta \sim \text{Normal}\,(0, 1)\\
&\text{Random effect (for } j \text{ subject)}\\
&\qquad \sigma_1^2 \sim \text{HalfCauchy}(5.)\\
&\qquad b_{\text{subject}_j} \sim \text{Normal}\,(0, \sigma_1^2)\\
&\text{Latent threshold}\\
&\qquad \mu = X\beta + b_{\text{subject}_j}\\
&\qquad \sigma_2^2 \sim \text{HalfCauchy}(2.5)\\
&\qquad \nu \sim \text{HalfNormal}(10.)\\
&\qquad \tau_{ij} \sim \text{Student-T}\,(\mu, \sigma_2^2, \nu)\\
&\text{Observed thresholds for each subject}\\
&\qquad \text{estimated\_mean}_{ij} \sim \text{Normal}\,(\tau_{ij}, \text{estimated\_std}_{ij})\\
&\text{Imputed model of censored threshold}\\
&\qquad \text{right\_censored} \sim \text{Student-T}\,(\mu_{k1}, \sigma_2^2, \nu)\\
&\qquad\qquad \textit{for } k_1 \text{ estimated\_mean}_{ij} > 1\\
&\qquad \text{left\_censored} \sim \text{Student-T}\,(\mu_{k2}, \sigma_2^2, \nu)\\
&\qquad\qquad \textit{for } k_2 \text{ estimated\_mean}_{ij} < 0
\end{aligned}
$$

For each individual, we modeled the estimated threshold of each expression and each task, along with the estimation uncertainty from the QUEST with a Bayesian Hierarchical Censored Model. The thresholds were modeled as a linear function of the full factorial of group (deaf signer, hearing nonsigner) * expression * task (intensity, signal), with the intercept of each subject as a random effect. This is equivalent to a linear mixed model with random intercept. Moreover, to account for the estimation error from the QUEST procedure, the estimated threshold is modeled as a realization of the latent (true) threshold with a standard deviation that is the same as the estimated standard deviation output from the QUEST procedure. Importantly, as explained above, because the presented intensity or signal information was limited to the range [0, 1], a threshold estimated above 1 is less reliable as it can never be directly observed. Thus, to account for these limitations, we added a penalization to the model log-likelihood using a censored variable representing the threshold estimated above 1 or below 0. We implemented an imputed censored model, where estimated values outside of [0, 1] are modeled as a set of additional random values that would have been censored. Thus, each censored observation introduces a random variable that would be added to the model log-likelihood.

A set of simpler alternative models were also tested: general linear model (no random effect, no latent parameters, and no additional censoring term), linear mixed model (no latent parameters and no censoring term), linear mixed model with latent parameters (but no censoring term), and the full model as described above. It is worth noting that the estimation was highly similar across all models, but the full model

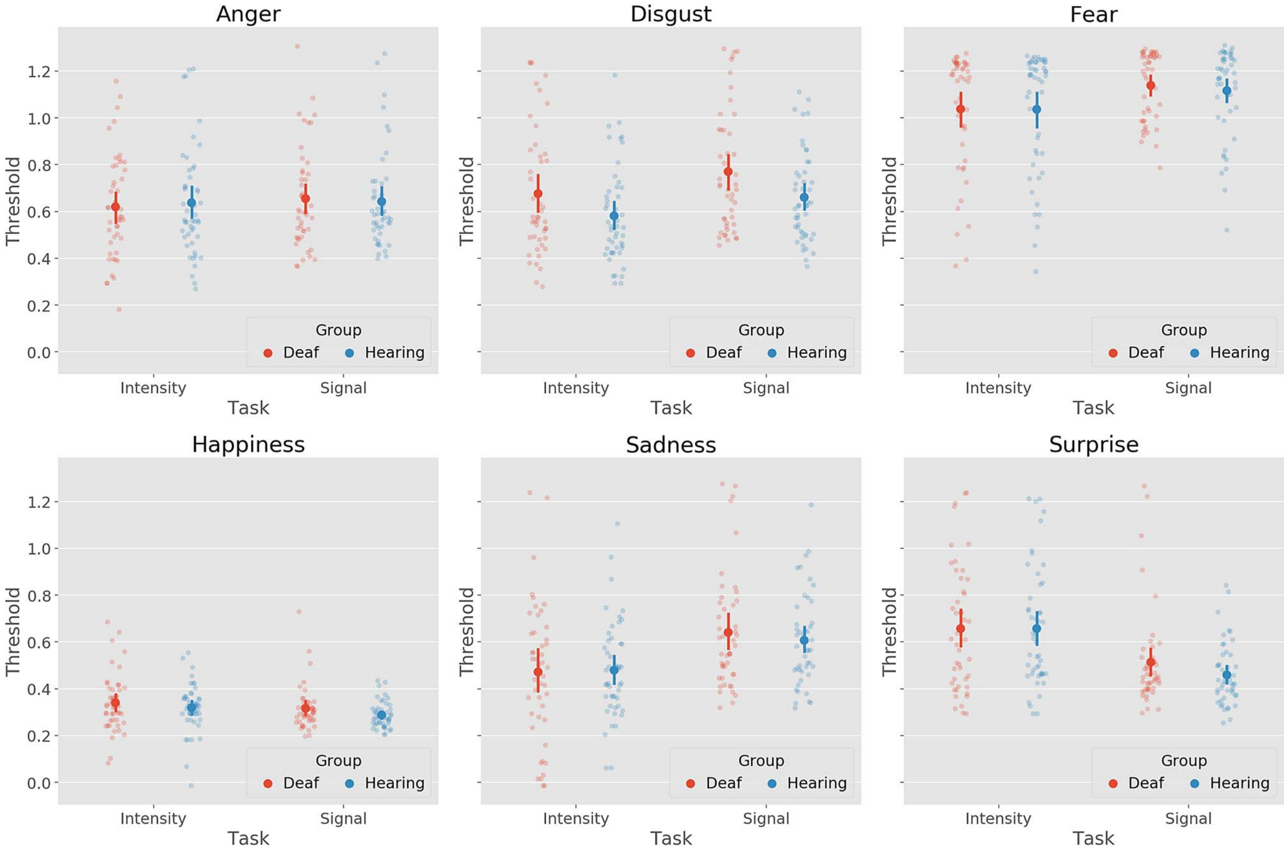

Figure 3 Descriptive statistics. Each subplot shows the threshold estimation for one facial expression of emotion. The scatter plots show the individual threshold estimations for one observer. Each group is color coded: red for deaf signers and blue for hearing nonsigners. The group mean is shown as a solid nontransparent dot on top of the scatter plot; the error bars show the 95% Bootstrap confidence interval of the mean.

yields the highest leave-one-out cross-validation score and the highest widely available information criterion (Vehtari et al., 2015; Spiegelhalter, Best, Carlin, & van der Linde, 2002). All the tested models and model comparisons are shown in the supplementary notebook.

The probabilistic model was built using PyMC3 and we sampled from the posterior distribution using NUTS. We ran four MCMC chains with 3,000 samples each; the first 1,000 samples were used for tuning the mass matrix and step size for NUTS and were discarded following this. Model convergence was diagnosed by computing Gelman and Rubin's (1992) convergence diagnostic (R-hat), examining the effective sample size, inspecting the mixing of the traces, and checking whether there is any divergent sample that has been returned from the sampler. Inferences were performed by computing directly on the posterior samples. Here, the posterior samples are a representation of the probability distribution of the unknown model parameters. We then used the posterior samples for (1) estimating facial expression recognition ability in deaf and hearing observers and (2) hypothesis testing of the group differences between deaf and hearing participants (i.e. computing the posterior differences and comparing these with 0).

Results

The descriptive results are shown in Figure 3. Both groups of participants showed similar recognition threshold means for each of the emotions. For example, happy was the easiest to recognize, as the estimated threshold was the lowest across all expressions. The most difficult expression was fear, with an estimated threshold consistently greater than 1. The mean recognition threshold for each group and each expression can be found in Supplementary Table 1.

Posterior distribution and estimation of the parameters of the Bayesian Hierarchical Censored Model are reported in Supplementary Figure 1 and Supplementary Table 2. The estimated threshold (i.e. overall intercept) was .642 [.576, .706], the values in brackets refer to the 95% highest probability density interval (95% HPD). At the group level, the deaf signers require a similar level of intensity/signal overall compared to hearing nonsigners, as the offset compared to the hearing nonsigners is estimated at .002 [−.090, .090]. There are nonzero estimates of the group * expression * task interaction terms from the linear equation part of the model, which we further quantified by computing the posterior conditional mean and performing a group comparison within each expression.

The posterior distribution of the latent threshold for each group, task, and expression is shown in Figure 4 and Supplementary Table 3. The two groups of observers do not show the same threshold estimation within each task and expression, as the posterior distributions are not completely overlapping with each other. To demonstrate this difference, we computed the posterior contrast between deaf and hearing observers for each expression and task (Figure 5). We report the contrast of deaf minus hearing participants. Interestingly, we found that deaf observers require more intensity and signal than the hearing

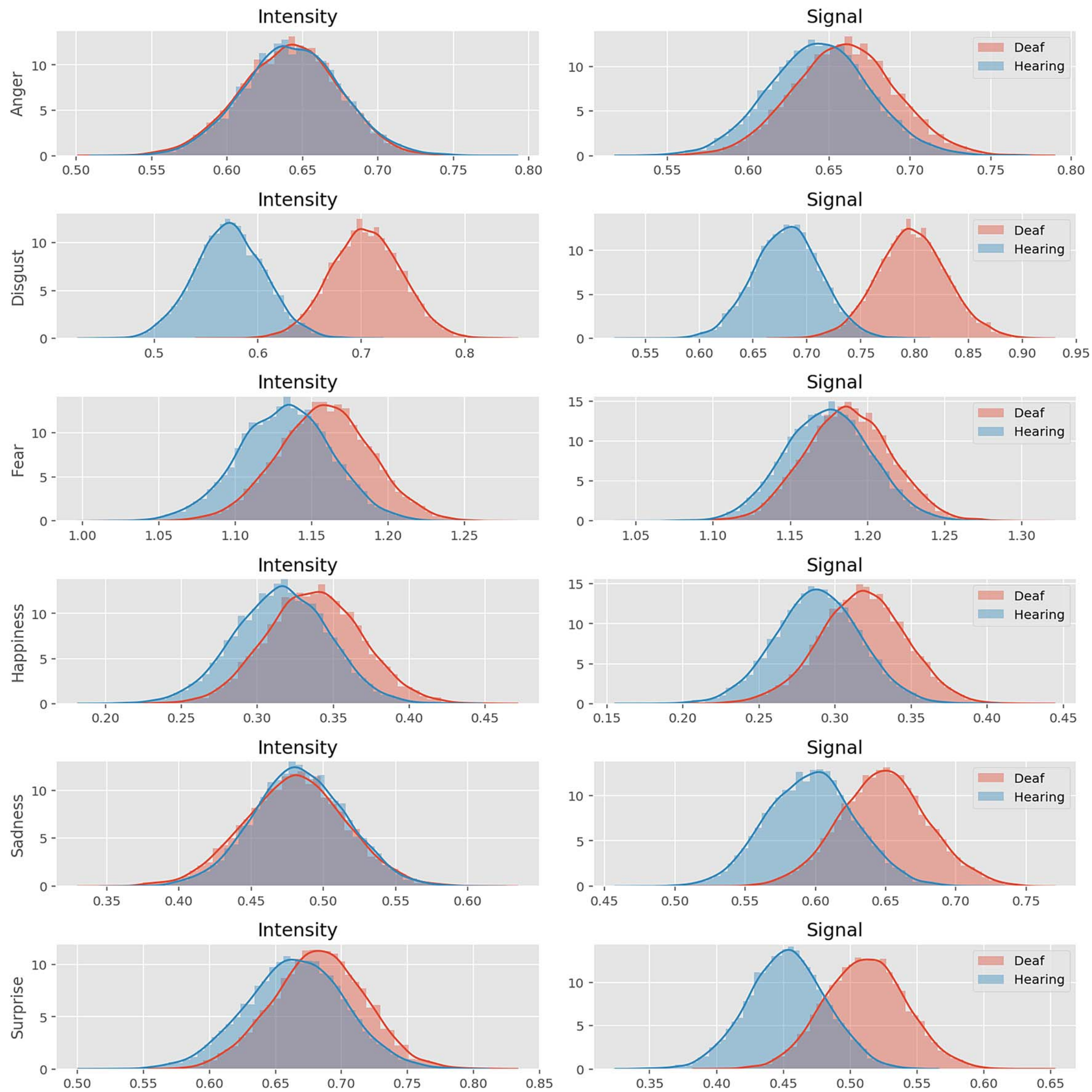

Figure 4 Posterior distribution of the thresholds for 75% recognition performance for each expression and task. Groups are color coded: red for deaf signers and blue for hearing nonsigners.

group to recognize the expression of *Disgust* accurately. The deaf observers need .130 [.039, .219] more intensity and .115 [.026, .198] more signal than the hearing participants. For the other expressions, the estimated thresholds for both tasks are similar for both groups of observers. Deaf observers also show a small increase of signal threshold for *Sadness* (.051 [−.038, .136]) and *Surprise* (.056 [−.026, .137]). However, as these latter estimated differences overlap with zero, more information is needed to draw conclusions about these conditions.

Discussion

The aim of this study was to assess facial expression recognition in deaf signers and hearing nonsigners using a psychophysical method to measure the quantity of intensity and signal needed to recognize an emotional expression. The sensitivity of this method combined with a large sample enabled us to obtain interesting and reliable observations about facial expression recognition in deaf signers.

Overall, our results suggest that deaf signers' facial expression recognition does not differ from hearing nonsigners in either the quantity of signal or intensity required to recognize the basic expressions, with the exception of disgust. For both deaf and hearing participants, the expression of happiness had the lowest signal and intensity thresholds and was therefore the easiest expression to recognize, whereas fear had the highest thresholds. The same pattern of results for happiness and fear has been reported in previous developmental studies using the same paradigm, which also tested hearing adults (Rodger et al., 2015, 2018). Further studies of emotion recognition in deaf children are necessary to further understand the developmental trajectory of emotion processing in the deaf population, and more specifically potential reasons for the difference found here in disgust recognition between deaf and hearing groups. Several studies report poorer performances in deaf children compared to age-matched hearing children in the labeling and categorization of facial expressions of emotion (e.g. Dyck, Farrugia, Shochet, & Homes-Brown, 2004; Ludlow, Heaton, Rosset, & Hills, 2010;

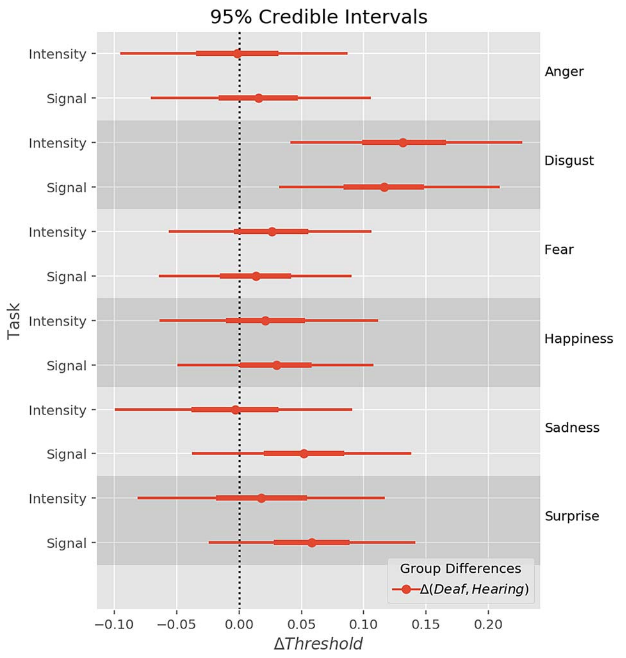

Figure 5 Forest plot of the posterior contrast of the thresholds for each expression and task between deaf signers and hearing nonsigners.

Most & Michaelis, 2012). These differences have been observed in young children from 6 years of age up to 18 years of age. However, other studies report no significant or subtle differences between deaf and hearing children in similar tasks involving facial emotion processing (e.g. Hosie, Gray, Russell, Scott, & Hunter, 1998; Jones et al., 2017; Ziv, Most, & Cohen, 2013). The heterogeneity of these results could be linked to different factors related to the study design, or to the participants characteristics such as age, use of hearing aids, language use (e.g. use of manual or oral communication), or age of first language exposure. They also suggest that more research is necessary to better understand the developmental trajectory of facial emotion processing in deaf children. Moreover, studies including both children and adults in the same experimental paradigm would be beneficial. In our study, 12 participants (about 25% of the sample) had a cochlear implant. It is possible that this artificial auditory input could have impacted the development of facial expression processing tested here. Indeed, these participants did not grow up with a total absence of auditory input. They could therefore link voice tone to facial expressions and processing could possibly have developed differently in comparison to their deaf counterparts with no cochlear implants. It has been shown that children with cochlear implants perform less well than typical hearing children and their performance during the recognition of emotional expressions is positively correlated with their linguistic skills (Wiefferink, Rieffe, Ketelaar, De Raeve, & Frijns, 2013). Further studies are required to directly compare facial expression processing in deaf with long-term use of cochlear implants and deaf with no access to auditory input. Here, we are confident that the statistical analysis computed in our study is sufficiently robust to ensure the reliability of our results.

Disgust was the only expression showing differences in recognition thresholds between the deaf signing and hearing nonsigning participants for both the signal and intensity conditions. Deaf signers had higher recognition thresholds than hearing nonsigners for both conditions. Therefore, deaf signers needed more facial information than their hearing counterparts to recognize disgust in comparison to the other facial expressions of emotion in order to maintain accuracy at a level of 75%. Atypical performance for disgust recognition was similarly reported in deaf children in a recent study of deaf and hearing children's recognition of the six basic emotions (Jones et al., 2017). In this study both deaf and hearing children showed the poorest performance for disgust and fear recognition overall, however deaf children showed poorer performance for disgust recognition compared to hearing children for both moving faces and faces that varied in emotional intensity. Jones et al. (2017) posited that deaf children's concepts of disgust in comparison to hearing children's may be less developed as they have been less exposed to conversations about emotions that may impact emotion recognition ability. However, it is not clear why the recognition of disgust would be uniquely affected within this context. Another possibility for this atypical performance in disgust recognition is that deaf people need more signal or intensity information because disgust may overlap with representations of other facial expressions like confusion, frustration, or uncertainty. It is possible that deaf people experience such expressions more frequently, so their perceptual categories for disgust and other expressions that are similar to them may be different to those of hearing people. Deafness can be considered as a culture. Indeed, Article 30 of the Convention on the Rights of Persons with Disabilities from the United Nations states that such persons are entitled "to recognition and support of their specific cultural and linguistic identity, including sign languages and deaf culture." By considering deafness as a cultural condition, the differences found in disgust processing can be attributed to specifities of this culture. In this way, perspectives on deafness are shifted from traditional deficit models of hearing loss toward one that sees deafness as representing one aspect of human diversity. As described in the introduction, the role of culture in visual cognition is timely and topical; our findings with deaf signers can be considered as new evidence that culture affects facial expression processing.

Performance for disgust recognition in another developmental study of deaf and hearing children was again lowest across all six basic expressions for both deaf and hearing children (Hosie et al., 1998). However, accuracy for disgust was higher for deaf children compared to hearing children. Nonetheless, this result is not a direct contradiction to our result with adult observers, as higher thresholds in our study were necessary to maintain a high level of performance (75% accuracy). Therefore, despite deaf children showing better scores than hearing children, accuracy overall was low in comparison to the 75% threshold in our study. High thresholds and accuracy are therefore not equivalent. In order to further address the question of atypical performance for disgust within deaf populations, a full parametric design is necessary to map out the complete psychometric function. Although it is clear that the facial expression of disgust is special in deaf signers (both adults and children), the reason for the atypical recognition performance remains unclear. Future studies are necessary to establish whether this difference relates to deafness or sign language experience. Further, now that atypical recognition of disgust in the deaf population has been detected, clearly there is a need for training in disgust recognition. Emotion recognition training in typically developing children has been shown to improve recognition performance (Pollux, Hall, & Guo, 2014). One pilot study with deaf children showed training improvements in emotion understanding but not facial expression recognition, so further studies in this area are clearly necessary (Dyck & Denver, 2003).

The absence of difference in recognition thresholds between deaf and hearing participants does not necessarily mean that the special status of faces for the deaf population has no impact on the decoding of facial expressions. Indeed, the current study only explored the coding of emotional facial expressions and not the other types of facial expressions. Because linguistic and emotional expressions are processed differently in deaf signers (Corina et al., 1999; McCullough et al., 2005), it is possible that only facial expressions used in sign language (i.e. linguistic facial expressions) are processed differently in deaf signers. It would therefore be necessary to extend research on facial expressions with a larger variety of expressions, in particular with both emotional and linguistic expressions in both hearing and deaf signers. In addition, it would also be interesting to clarify whether deaf adults would benefit from the presentation of dynamic emotional expressions over static images. We very recently showed that the performance for the recognition of *static* as compared to *dynamic* expressions was notably less effective in fragile face processing systems, such as brain-damaged patients (Richoz, Jack, Garrod, Schyns, & Caldara, 2015) and elderly adults (Richoz, Lao, Pascalis, & Caldara, 2018) by using a database of stimuli controlled for the amount of *low-level* information carried over time (Gold et al., 2013). As sign language is dynamic, we could expect similar results in the deaf population, a question that needs to be addressed in a future study.

In the current study, we aimed to provide a benchmark for the recognition of facial expressions of emotion in deaf signers, using young adult hearing participants as a baseline. Thus, instead of performing a conventional series of *t*-tests for all possible combinations of each groups' expression recognition measures corrected for multiple comparisons, we decided to improve the estimation of accuracy for each condition by properly accounting for measurement errors using weakly informative priors. A weakly informative prior is implicit in comparison to a default uniform prior. It gives slightly more prior weight in some region of the domain of the parameters. Thus, if there is very little data, a weakly informative prior will strongly influence the posterior inference. However, with a reasonably large amount of data (as is the current case), the likelihood will dominate the posterior inference and the prior will not be important. Moreover, a weakly informative prior greatly improves model convergence. More information regarding prior choice can be found in https://github.com/stan-dev/stan/wiki/Prior-Choice-Recommendations. With this novel threshold paradigm, we take advantage of the uncertainty of the individual threshold estimation output by the adaptive maximum likelihood procedure. By using a hierarchical mixed-effect model, we account for the individual differences and provide a more precise group estimation. Moreover, we account for the bias in the threshold estimations that are higher than one with a censoring likelihood as a penalty. We hope that future studies with similar paradigms can take advantage of the model estimation we provide here, to construct more informative models.

To conclude, this study offers new insights into the coding of facial expressions of emotion in deaf signers. Despite the central importance of facial expressions in deaf communication, overall we observe similar intensity and signal thresholds for both deaf signers and hearing nonsigners for facial expression recognition. Further studies are necessary to examine potential differences in the recognition of both emotional and linguistic expressions, in both hearing and deaf signers. Exceptionally, recognition performance for disgust was poorer for deaf signers compared to hearing nonsigners. The atypical processing of disgust expressions has similarly been reported in deaf children. Future studies should try to establish how the experience of deafness and sign language may interfere in the construction and understanding of expressions of disgust. One consideration is that the Action Units for sign language may overlap with the Action Units for emotions, but not uniformly. That is, the Action Units for disgust may incorporate articulations that are more similar to those employed in sign language than is the case for other emotional expressions. Our results provide a reliable benchmark for the intensity and signal thresholds used for expression recognition in young deaf adults.

Funding

This study was supported by the Swiss National Science Foundation (n° 100014_156490/1) awarded to Roberto Caldara.

Conflict of Interest

No conflicts of interest were reported.

References

Bayet, L., Quinn, P. C., Laboissière, R., Caldara, R., Lee, K., & Pascalis, O. (2017). Fearful but not happy expressions boost face detection in human infants. *Proceedings of the Royal Society of London B*, *284*, 1–9. doi:10.1098/rspb.2017.1054

Brentari, D., & Crossley, L. (2002). Prosody on the hands and face: Evidence from American Sign Language. *Sign Language & Linguistics*, 5, 105–130. doi:10.1075/sll.5.2.03bre

Caldara, R. (2017). Culture reveals a flexible system for face processing. *Current Directions in Psychological Science*, *26*, 249–255. doi:10.1177/0963721417710036

Campanella, S., & Belin, P. (2007). Integrating face and voice in person perception. *Trends in Cognitive Sciences*, *11*, 535–543. doi:10.1007/978-1-4614-3585-3

Campbell, R., Woll, B., Benson, P. J., & Wallace, S. B. (1999). Categorical perception of face actions: Their role in sign language and in communicative facial displays. *The Quarterly Journal of Experimental Psychology*, 52, 67–95. doi:10.1080/713755802

Carton, J. S., Kessler, E. A., & Pape, C. L. (1999). Nonverbal decoding skills and relationship well-being in adults. *Journal of Nonverbal Behavior*, *23*, 1–100. doi:10.1023/A:1021339410262

Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., & Lepore, F. (2008). Audio-visual integration of emotion expression. *Brain Research*, *1242*, 126–135. doi:10.1016/j.brainres.2008.04.023

Corina, D. P., Bellugi, U., & Reilly, J. (1999). Neuropsychological studies of linguistic and affective facial expressions in deaf signers. *Language and Speech*, *2–3*, 307–331. doi:10.1177/00238309990420020801

Crivelli, C., Russell, J. A., Jarillo, S., & Fernández-Dols, J. M. (2016). The fear gasping face as a threat display in a Melanesian society. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 12403–12407. doi:10.1073/pnas.1611622113

Lewkowicz, D. J.(2014). Early experience and multisensory perceptual narrowing. *Developmental Psychobiology*, 56, 292–315. doi:10.1002/dev.21197

Dyck, M., & Denver, E. (2003). Can the emotion recognition ability of deaf children be enhanced? A pilot study. *Journal of Deaf Studies and Deaf Education*, 8, 348–356. doi:10.1093/deafed/eng019

Dyck, M. J., Farrugia, C., Shochet, I. M., & Holmes-Brown, M. (2004). Emotion recognition/understanding ability in hearing or vision-impaired children: Do sounds sights, or words make the difference? *Journal of Child Psychology and Psychiatry*, *45*, 789–800. doi:10.1111/j.1469-7610.2004.00272.x

Feldman, R. S., Philippot, P., & Custrini, R. J. (1991). Social competence and nonverbal behavior. In *Fundamentals of nonverbal behavior* (pp. 329–350). New York, NY: Cambridge University Press.

Geangu, E., Ichikawa, H., Lao, J., Kanazawa, S., Yamaguchi, M. K., Caldara, R., & Turati, C. (2016). Culture shapes 7-month-olds perceptual strategies in discriminating facial expressions of emotion. *Current Biology*, *26*, R663–R664. doi:10.1016/j.cub.2016.05.072

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*, 457–511. doi:10.1214/ss/1177011136

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014). Perceptions of emotion from facial expressions are not culturally universal: Evidence from a remote culture. *Emotion*, 14, 251–262. http://dx.doi.org/10.1037/a0036052

Gold, J. M., Barker, J. D., Barr, S., Bittner, J. L., Bromfield, W. D., Chu, N., & Srinath, A. (2013). The efficiency of dynamic and static facial expression recognition. *Journal of Vision*, *13*, 1–12. doi:10.1167/13.5.23

Grossman, R. B., & Kegl, J. (2007). Moving faces: Categorization of dynamic facial expressions in American Sign Language by deaf and hearing participants. *Journal of Nonverbal Behavior*, *31*, 23–38. doi:10.1007/s10919-006-0022-2

Hall, E. (1966). *The hidden dimension* (). Garden City, NY: Doubleday.

Hosie, J. A., Gray, C. D., Russell, P. A., Scott, C., & Hunter, N. (1998). The matching of facial expressions by deaf and hearing children and their production and comprehension of emotional labels. *Motivation and Emotion*, *22*, 293–313. doi:10.1023/A:1021352323157

Izard, C., Fine, S., Schultz, D., Mostow, A., Ackerman, B., & Youngstrom, E. (2001). Emotion knowledge as a predictor of social behavior and academic competence in children at risk. *Psychological Science*, *12*, 18–23. doi:10.1111/1467-9280.00304

Jack, R., Caldara, R., & Schyns, P. (2012a). Internal representations reveal cultural diversity in expectations of facial expressions of emotions. *Journal of Experimental Psychology: General*, *141*, 19–25. doi:10.1037/a0023463

Jack, R., Garrod, O., Yu, H., Caldara, R., & Schyns, P. G. (2012b). Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences*, *109*, 7241–7244. doi:10.1073/pnas.1200155109

Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. *Current Biology*, *19*, 1543–1548. doi:10.1016/j.cub.2009.07.051

Jack, R. E., & Schyns, P. G. (2015). The human face as a dynamic tool for social communication. *Current Biology*, *25*, R621–R634. doi:10.1016/j.cub.2015.05.052

Jack, R. E., Sun, W., Delis, I., Garrod, O. G. B., & Schyns, P. G. (2016). Four not six: Revealing culturally common facial expressions of emotion. *Journal of Experimental Psychology: General*, *145*, 708–730. doi:10.1037/xge0000162

Jones, A. C., Gutierrez, R., & Ludlow, A. K. (2017). The role of motion and intensity in deaf children's recognition of real human facial expressions of emotion. *Cognition and Emotion*, *9931*, 1–14. doi:10.1080/02699931.2017.1289894

Jones, M. A. (2002). Deafness as culture: A psychosocial perspective. *Disability Studies Quarterly*, *22*, 51–60. doi:10.18061/dsq.v22i2.344

King-Smith, P. E., Grigsby, S. S., Vingrys, A. J., Benes, S. C., & Supowit, A. (1994). Efficient and unbiased modifications of the QUEST threshold method: Theory, simulations, experimental evaluation and practical implementation. *Vision Research*, *34*, 885–912. doi:10.1016/0042-6989(94)90039-6

Lewkowicz, D. J., & Ghazanfar, A. A. (2009). The emergence of multisensory systems through perceptual narrowing. *Trends in Cognitive Sciences*, *13*, 470–478. doi:10.1016/j.tics.2009.08.004

Liddell, S. K. (2003). *Grammar, gesture, and meaning in American Sign Language*. New York, NY: Cambridge University Press.

Ludlow, A., Heaton, P., Rosset, D., Hills, P., & Deruelle, C. (2010). Emotion recognition in children with profound and severe deafness: Do they have a deficit in perceptual processing ? *Journal of Clinical and Experimental Neuropsychology*, *32*, 923–928. doi:10.1016/j.cogbrainres.2004.08.012

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska Directed Emotional Faces - KDEF, CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, ISBN 91-630-7164-9.

McCullough, S., Emmorey, K., & Sereno, M. (2005). Neural organization for recognition of grammatical and emotional facial expressions in deaf ASL signers and hearing nonsigners. *Cognitive Brain Research*, *22*, 193–203. doi:10.1016/j.cogbrainres.2004.08.012

Most, T., & Michaelis, H. (2012). Auditory, visual, and auditory–visual perceptions of emotions by young children with hearing loss versus children with normal hearing. *Journal of Speech Language and Hearing Research*, *55*, 1148–1162. doi:10.1023/A:1021352323157

Nowicki, S. Jr., & Duke, M. P. (1992). The association of children's nonverbal decoding abilities with their popularity, locus of control, and academic achievement. *The Journal of Genetic Psychology*, *153*, 385–394. doi:10.1080/00221325.1992.10753734

Pollux, P., Hall, S., & Guo, K. (2014). Facial expression training optimises viewing strategy in children and adults. *PLoS One*, *9*, e105418. doi:10.1371/journal.pone.0105418

Reilly, J. S., & Bellugi, U. (1996). Competition on the face: Affect and language in ASL motherese. *Journal of Child Language*, *23*, 219–239. doi:10.1017/S0305000900010163

Reilly, J., Mcintire, M. L., & Seago, H. (1992). Affective prosody in American Sign Language. *Sign Language Studies*, *75*, 113–128. doi:10.1353/sls.1992.0035

Richoz, A.-R., Jack, R. E., Garrod, O. G., Schyns, P. G., & Caldara, R. (2015). Reconstructing dynamic mental models of facial expressions in prosopagnosia reveals distinct representations for identity and expression. *Cortex*, *65*, 50–64. doi:10.1016/j.cortex.2014.11.015

Richoz, A. R., Lao, J., Pascalis, O., & Caldara, R. (2018). Tracking the recognition of static and dynamic facial expressions of emotion across the life span. *Journal of Vision*, *18*, 1–27. doi:10.1167/18.9.5

Rodger, H., Lao, J., & Caldara, R. (2018). Quantifying facial expression signal and intensity use during development. *Journal of Experimental Child Psychology*, *174*, 41–59. doi:10.1016/j.jecp.2018.05.005

Rodger, H., Vizioli, L., Ouyang, X., & Caldara, R. (2015). Mapping the development of facial expression recognition. *Developmental Science*, *18*, 926–993. doi:10.1111/desc.12281

Sims, J. A., & Pelli, D. G. (1987). The ideal psychometric procedure. *Investigative Ophthalmology & Visual Science*, *28*, 366.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B*, *64*, 583–639. doi:10.1111/1467-9868.00353

Tiddeman, B., Burt, M., & Perrett, D. (2001). Prototyping and transforming facial textures for perception research. *IEEE Computer Graphics and Applications*, *21*, 42–50. doi:10.1109/38.946630

Turano, M. T., Lao, J., Richoz, A.-R., de Lissa, P., Degosciu, S. B. A., Viggiano, M. P., & Caldara, R. (2017). Fear boosts the early neural coding of faces. *Social Cognitive and Affective Neuroscience*, *12*, 1959–1971. doi:10.1093/scan/nsx110

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*, 1413–1432. doi:10.1007/s11222-016-9696-4

Watanabe, K., Matsuda, T., Nishioka, T., & Namatame, M. (2011). Eye gaze during observation of static faces in deaf people. *PLoS One*, *6*, e16919. doi:10.1371/journal.pone.0016919

Watson, A., & Pelli, D. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, *33*, 113–120. doi:10.3758/BF03202828

Wiefferink, C. H., Rieffe, C., Ketelaar, L., De Raeve, L., & Frijns, J. H. (2013). Emotion understanding in deafchildren with a cochlear implant. *Journal of Deaf Studies and Deaf Education*, 18, 175–186. doi:10.1093/deafed/ens042

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods*, *42*, 671–684. doi:10.3758/BRM.42.3.671

Ziv, M., Most, T., & Cohen, S. (2013). Understanding of emotions and false beliefs among children versus deaf children. *Journal of Deaf Studies and Deaf Education*, 18, 161–174. doi:10.1093/deafed/ens073