

RESEARCH ARTICLE

The age bias in labeling facial expressions in children: Effects of intensity and expression

Dafni Surian^{1*}, Carlijn van den Boomen²

1 Department of Developmental Psychology, Utrecht University, Utrecht, The Netherlands, **2** Department of Experimental Psychology, Helmholtz Institute, Utrecht University, Utrecht, The Netherlands

* dafni.surian@gmail.com

Abstract

Emotion reasoning, including labeling of facial expressions, is an important building block for a child's social development. This study investigated age biases in labeling facial expressions in children and adults, focusing on the influence of intensity and expression on age bias. Children (5 to 14 years old; N = 152) and adults (19 to 25 years old; N = 30) labeled happiness, disgust or sadness at five intensity levels (0%; 25%; 50%; 75%; and 100%) in facial images of children and adults. Sensitivity was computed for each of the expression-intensity combinations, separately for the child and adult faces. Results show that children and adults have an age bias at low levels of intensity (25%). In the case of sadness, children have an age bias for all intensities. Thus, the impact of the age of the face seems largest for expressions which might be most difficult to recognise. Moreover, both adults and children label most expressions best in adult rather than child faces, leading to an other-age bias in children and an own-age bias in adults. Overall, these findings reveal that both children and adults exhibit an age bias in labeling subtle facial expressions of emotions.

OPEN ACCESS

Citation: Surian D, van den Boomen C (2022) The age bias in labeling facial expressions in children: Effects of intensity and expression. PLoS ONE 17(12): e0278483. <https://doi.org/10.1371/journal.pone.0278483>

Editor: Peter A. Bos, Leiden University, NETHERLANDS

Received: May 2, 2022

Accepted: November 17, 2022

Published: December 2, 2022

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0278483>

Copyright: © 2022 Surian, van den Boomen. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The data is in the possession of author Carlijn van den Boomen. There are ethical restrictions that prohibit sharing the data set of this manuscript. Participants and/or their caregivers did not consent to make data

Introduction

The ability to perceive facial expressions of emotion (to use the traditional reference) is an important building block of social and emotional development [1–3]. This ability develops throughout childhood: infants start to discriminate and differentially process different facial configurations at four to seven months of age [4–7], and the labeling of facial expressions of emotion refines until 10 years of age [8], or even longer [9]. However, in recent years it is increasingly recognised that what has been traditionally labeled as 'facial expressions of emotion' is based on several assumptions [10–12], particularly that the displayed facial configuration as a result of muscle movement reflects the emotional state of the actor. In addition, the ability to 'recognise' these configurations is now understood to include both the visual processing of the configuration as the understanding of the emotional state, and to rely on a wide range of processes [12]. Furthermore, in daily life an understanding of someone's emotional state is not only based on the face (although the face plays a crucial role from early on in life [3]) but also on other signals from the actor and the context [12–14]. Here we pose that the study of emotion reasoning might be even more complex as the visual processing of the configurations seems to depend on specific characteristics of the face itself. The current study

publicly available. Consequently, the Faculty Ethical Research Board of the Faculty of Social and Behavioural Sciences at Utrecht University restricts data sharing. Nevertheless, researchers aiming to replicate the study's findings could request access to the data via this board. The ethical board can be reached via the secretary Mr. J. Tenkink- de Jong at j.f.tenkink-dejong@uu.nl.

Funding: The authors received no specific funding for this work.

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

explores which combinations of facial characteristics—specifically age, expression, and intensity—should be taken into account in future studies on emotion reasoning in children.

One of the stimulus characteristics known to influence emotion reasoning in adults is the age of the face on which the emotion appears. Previous research found an own-age bias in adults, which means that adults are better at labeling emotions in faces of adults of the same age group than in older or younger faces [15]. A 50-year-old person will for example be better at labeling the emotions of another middle-aged person than those of an elderly. Although an own-age bias for face recognition is consistently reported [16], only a couple of studies have investigated own-age bias for labeling of facial expressions in children. Some did not find an own-age bias in children between the age of 5 and 17 years [17, 18]: the children performed as good as the adults on all three age groups of the shown faces. Note however that Griffiths and colleagues [17] report that both children and adults label happy, sad and angry expressions more accurately in child than in adult faces, and disgust more accurately in adult faces. However, they do not interpret this as an age-bias. An own-age bias was revealed by one study in adolescents (11 to 14 years [19]). In this study, it should be noted that it cannot be excluded that presentation order (all adult faces presented before all child faces) affected the reported bias in the latter study. As such, to date the existence of an own-age bias for emotion recognition in children remains inconclusive.

The age bias could interact with the intensity of the facial expression. As highlighted by Ruba and Pollak [12], facial stimuli in experiments often display full-intensity facial configurations, which are infrequently present in human interactions. Emotion reasoning seems to be more difficult for subtle facial configurations: Gao and Maurer [9] created twenty intensities with increments of 5%, such as 5% happy, 10% happy, etc., until 100% happy. Children could label happiness similarly well as adults already at an age of 5 years old (youngest age tested), not only when an extreme (100%) display of happiness was shown, but also at more subtle intensities. However, children needed a more intense face to label fearful and sad faces than adults did. Yet, Gao and Maurer [9] only used adults' faces in their research. As such, it is unclear how children would perform on a task requiring them to label subtle expressions in faces of children. This was investigated by Griffiths and colleagues [17] who reported no interaction between intensity and face age. However, they presented only two intensities (i.e. 'original' and 'caricatured') with an undefined specific intensity level. Based on observation of the published images, both intensities were high compared to the lowest intensities in the research by Gao and Maurer [9]. In their next study on own-age bias, eight intensities were included [20], but not analysed due to the focus of that study. Thus, although there seems to be no own-age bias for highly intense emotions, it is unknown whether this exists for more subtle emotions.

Finally, the interaction between the age bias and the intensity could be further complicated by that it might depend on the facial configuration related to specific emotional labels. This is due to the finding that emotion reasoning seems to develop at different paces for separate expressions. As reviewed by Herba and Phillips [21], labeling of facial expressions gradually improves throughout childhood. However, the rate of improvement differs between expressions, with happiness being labeled as accurately as adults at the youngest age (e.g. at 5 years; [8]). While several studies report sadness and anger to be labeled next, followed by surprise and fear [21], others report different orders [8, 22]. However, regardless of the specific order of expressions that are labeled similar to adults and the need to unravel the multiple underlying cognitive processes that affect the development [12], there seems to be consensus that children's ability to label expressions depend on the expression itself.

Overall, previous research investigated the effect of several stimulus characteristics on emotion reasoning separately. However, none has combined the age of the face, the expression,

and the intensity of the expression, and thus have not explored the complex interplay between these characteristics. The aim of the current study is to get a better understanding of these combined characteristics. We focus on the age of the face, as this characteristic is the least well understood. As such, we investigate whether there is an age bias for labeling facial expressions in typically developing children and take into account the intensity at which an expression is shown, and the type of expression. This study combines the stimulus presentations used by Griffiths and colleagues [17, 20] and by Gao and Maurer [9]: children perceive images of faces of both adults and children, in which different expressions (happy, sad, disgust) are presented at different levels of intensity (0, 25, 50, 75 & 100 percent). For comparison, the task was also completed by adult participants. Due to the importance of emotion reasoning for social interaction [23, 24], this knowledge would help stimulate social interactions with typically developing children. Moreover, even though atypically developing children might benefit from a different combination of characteristics [25], the current findings could provide a starting point for optimizing training programs in these populations as well [26].

The original hypothesis posed that the age of the face, the intensity, the emotional label, and the age of the participant would interactively affect sensitivity to an emotion. However, this hypothesis could not be tested, because the data was extremely skewed and as such not normally distributed. Therefore, we needed to use non-parametric statistics, that do not allow interaction-analyses. Therefore, a more limited set of hypotheses were posed, with a focus on the age bias. For emotions with a high intensity, there is no clear direction in the hypothesis: while there is an indication that children are better at recognizing emotions in faces of children [19], others find that the age of the face does not affect children's performance [17, 20]. For lower intensities, particularly at 25%, it is expected that there will be an age-bias, because the sensitivity is likely to be lower [9] and therefore more affected by other stimulus characteristics.

Methods

Participants

One hundred fifty-two children and 30 adults participated in the study. In Table 1 the distribution of the participants across age groups and gender can be found. The difference in gender for the children and adults is not significant (chi-square = 2.079; $p = .15$). Note that although the sample size differs between the children and adults, both samples yield large power using the current experimental set-up [27]. All participants had normal or corrected to normal vision and had no diagnosis of a psychiatric illness, except that in the group of children, three had a diagnosis ADHD, four a diagnosis Autism Spectrum Disorder, and one both these disorders. Removing these children from the analyses did not affect the conclusions, and thus these children were included in the final sample. Thirteen additional children were excluded from the analysis. Four children did not complete the task due to lack of motivation and nine children could not complete the task due to a technical error.

Table 1. Distribution of the participants across age and gender.

	<i>Children</i>	<i>Adults</i>	<i>Total</i>
<i>Number</i>	152	30	182
<i>Age in years</i>	5,9 to 14,6 ($M = 10$, $SD = 1,9$)	19,5 to 25,3 ($M = 21,7$, $SD = 1,2$)	5,9 to 25,3 ($M = 11,94$, $SD = 4,68$)
<i>Female</i>	79 (52%)	20 (66,6%)	99 (54,4%)
<i>Male</i>	72 (48%)	10 (33,3%)	82 (45,6%)

<https://doi.org/10.1371/journal.pone.0278483.t001>

The adult participants were recruited at Utrecht University. Most of them were students of the bachelor program in Psychology, and received study credits as compensation for their participation. The children were recruited among the visitors of a science museum. All parents gave written informed consent for their children's participation in the study. Children above the age of 12 and the adult participants gave written informed consent themselves. The children received a certificate and a yo-yo for participating in the study. A local ethical committee of the Faculty of Behavioral Sciences at Utrecht University, The Netherlands, approved the experimental procedure. The study has been conducted following the guidelines of the Declaration of Helsinki (2008).

Stimuli

32 pictures were selected from the Radboud Faces Database (validated in adults [28] and children [29]). The 32 pictures were photographs of eight models: four children, of which two girls (number 64 and 65) and two boys (number 42 and 63), and four adults, of which two women (number 27 and 61) and two men (number 33 and 71). Although the age of the selected models is unknown, the age of all child models in the database is between 7 and 12 years, with the age of one child model being unknown (number 29). Each model posed with one happy, one sad, one disgusted and one neutral expression. The selection of types of expressions was based on the results of Gao and Maurer [9], who revealed that children find it particularly difficult to recognize sad faces, and confuse these with neutral or disgust. On the contrary, already at 5 years of age children could label happy expressions as well as adults, even for the lowest intensities, which is why we included this expression as a proof of principle for our experiment. We did not add additional expressions as a pilot study revealed that the experiment became too long for the participants when an additional expression was added. The pictures have a resolution of 1024×681 pixels. The selected photos are the ones in which adults categorize the expression with the highest consensus ($M = 88\%$; [28]). For each expression four levels of intensity were created: 25%, 50%, 75% and 100%. This was not done for the neutral expression, which represented the 0% intensity. Similar to Gao and Maurer [9] this was done using the program MorphX (<http://www.norrkross.com/software/morphx/MorphX.php>). Distortions resulting from the morphing process were fixed with Photoshop (version CC2014), and the background colour was changed into RGB 108x108x108. This created 104 stimuli (8 models x 3 emotions x 4 intensity levels + 8 neutral photos, 1 for each of the 4 models). The faces were resized to 11 x 16,7 degrees of visual angle at a viewing distance of 57 centimetre (measured from the eyes to the centre of the screen). The stimuli were displayed on an HP-laptop, the Elitebook 840G3, with an external keyboard.

Procedure

Testing the children took place in a quiet corner of the museum, illuminated by natural light. Three participants could be tested at the same time, each on a different laptop. The adult participants completed the study in a lab at the university building, with dimmed lights. In both situations, it was ensured that light was present but did not reflect on the screen. An external keyboard was provided for each laptop, to ensure the participants could easily reach the keys. We aimed to create a digital version of the set-up by Gao & Maurer [9]. Four stickers were placed on the four keys needed to select the chosen emotions in the task, to make them easily recognizable. Furthermore, a paper showing the key-emotion combinations was placed between the keyboard and the laptop. The paper served as a reminder for the combinations, but was small enough not to occlude the screen. The participants were instructed, by means of a short story to make it easier for the children to understand (see [S1 Appendix](#); copied and

adapted from Gao & Maurer [9]), to categorize the faces on the screen as neutral, happy, sad or disgusted using the keyboard. The relevant keys were z, x, n, m, with two sets of key-expression combinations that were randomized between participants.

The experiment started with eight practice trials including faces of all expressions and ages with 100% intensity. In these practice trials, a reminder with the key-label combinations was displayed after every answer. After each choice, the participants were reminded of which key corresponded to each expression with a picture appearing on the screen. After the practice trials, the actual task started, which consisted of 104 pictures, separated in three blocks of 35, 35 and 34 pictures. Per trial, a grey screen (RGB 108x108x108) appeared for a jittered time between 500 and 700 ms. After this, the face was presented, which remained on the screen until a choice between the four labels was made. After a response was provided, the participants saw on the screen that they had earned 1 point. This point was earned regardless of their choice, to avoid providing feedback on the correctness of the answer. The participants pushed the spacebar to continue to the next trial. At the end of every block, the participants saw that they just reached a new level and how many levels they had left. This was done to split the task into three blocks, to ensure the participants could have a break, and also to make the experiment into a game. The experiment lasted 15 to 20 minutes, including the explanation.

Analysis

To investigate labeling ability, we calculated the sensitivity of the participants to every combination of expression, intensity and face of an adult or of a child. We choose to compute sensitivity instead of the percentage correct responses, to correct answers for wrongly choosing the emotion. Because each specific combination of characteristics was presented in four trials per combination, several participants showed 100% hits and 0% false alarms or vice versa. As such, d' could not be computed. Therefore, we computed sensitivity by means of A' (aprime), using the following formula [30]:

$$A' = 0.5 + (\text{sign}(\text{HR} - \text{FAR}) * (((\text{HR} - \text{FAR})^2) + \text{abs}(\text{HR} - \text{FAR}))) / (4 * \max(\text{HR}, \text{FAR}) - (4 * \text{HR} * \text{FAR}));$$

HR stands for hit rate (the percentage of correctly categorized faces as displaying a specific emotion), FAR stands for false alarm rate (the percentage of faces wrongly categorized as displaying this emotion) and A' or aprime stands for the sensitivity of the participant to an emotion. Aprime ranges from 0 to 1, where 0.5 is chance level and 1 is maximum sensitivity (perfect score).

To investigate per age-group of the participants (measured between subjects) the impact of age of the face, intensity of the expression and expression (all measured within subjects) on sensitivity, we conducted non-parametric analyses, because of the extreme skewness and hence non-normal distribution of the data. As the focus of the hypotheses is on own-age bias, the difference in sensitivity between child and adult faces was tested with Wilcoxon Signed Ranks tests. This was done per expression and per intensity of the face, and separately for the child and adult participants. Performance of multiple comparisons was corrected by dividing the alpha value of .05 by the number of comparisons per age-group, leading to alpha 0.004.

Furthermore, as the age-range within the group of children was quite broad (5 to 14 years) we also ran exploratory analyses including age as a continuous variable to reach a more comprehensive understanding of the effect of age on sensitivity to expressions, with a focus on the most subtle expressions. As such, we computed non-parametric exploratory correlation analyses between age of the participant and 1) sensitivity to each of the different expressions with all intensities combined; 2) sensitivity to each of the different expressions at 25% intensity, based on the results of the planned analyses described above; and 3) direction of the bias (computed

as a prime of adult faces minus a prime of child faces) for each of the expressions at 25% intensity.

Results

To evaluate the effect of age of the face, combined with the intensity and expression, we tested if children and adults were more sensitive to emotions on the faces of children or adults. An overview of the results can be found in Table 2, and boxplots for sensitivity at 25% intensity in Fig 1. In children, the Wilcoxon Signed Ranks tests showed that for disgust at 25% and 50%, children have a higher sensitivity to the expression displayed on faces of children than on faces of adults: they have an own-age bias (25% intensity: $Z = -6.9$, $p < .001$; 50% intensity (bias direction based on boxplots): $Z = -4.1$, $p < .001$). For other expressions, this own-age bias was not found. Instead, other-age biases were found for some intensities: children appear to have a higher sensitivity to the expression on adult than child faces for neutral faces, happiness at an intensity of 25%, and sadness on all intensity levels (neutral: $Z = -3.3$, $p = .001$; happiness at 25% intensity: $Z = 5.2$, $p < .001$; sadness at 25% intensity: $Z = 8.5$, $p < .001$; sadness at 50% intensity: $Z = 5.8$, $p < .001$; sadness at 75%: $Z = 7.4$, $p < .001$; sadness at 100% intensity: $Z = 7.7$, $p < .001$). For all other intensities of the expressions there was no difference in sensitivity between the child and adult face (all $p > .004$), and thus no bias was found. In adults, for

Table 2. Overview of the results and medians of the hypothesis about age bias for the children and the adults. Note that for the median A' 0.5 represents guessing and 1 represents perfect performance.

Participant	Expression	Intensity	Median adult faces	Median child faces	Bias (p-value)	
Children	Neutral		0.92	0.9	Other-age bias (0.001)	
		Disgust	25%	0.5	0.81	Own-age bias ($< .001$)
			50%	1	1	Own-age bias ($< .001$)
			75%	1	1	No
			100%	1	1	No
	Sadness	25%	0.81	0.5	Other-age bias ($< .001$)	
		50%	0.93	0.88	Other-age bias ($< .001$)	
		75%	1	0.94	Other-age bias ($< .001$)	
		100%	1	0.94	Other-age bias ($< .001$)	
	Happiness	25%	0.88	0.81	Other-age bias ($< .001$)	
		50%	1	1	No	
		75%	1	1	No	
100%		1	1	No		
Adults	Neutral		0.93	0.92	No	
		Disgust	25%	0.81	0.81	No
			50%	1	1	No
			75%	1	1	No
	Sadness	25%	0.875	0.555	Own-age bias (0.001)	
		50%	0.94	0.94	No	
		75%	1	1	No	
		100%	1	1	No	
	Happiness	25%	0.825	0.81	Own-age bias (0.002)	
		50%	1	1	No	
		75%	1	1	No	
		100%	1	1	No	

<https://doi.org/10.1371/journal.pone.0278483.t002>

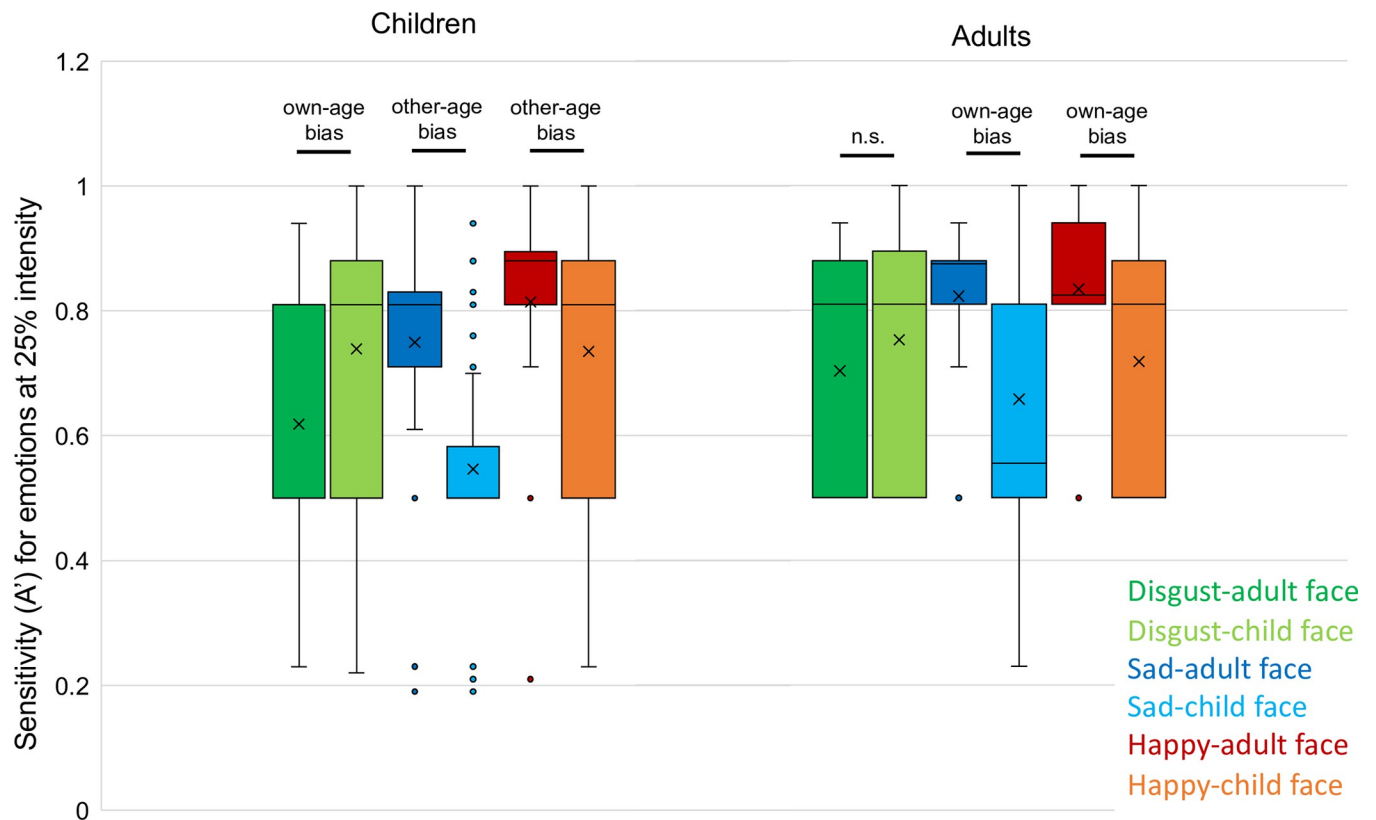


Fig 1. Boxplots of sensitivity of children (left) and adults (right) for expressions at 25% intensity, separately for the different combinations of displayed expressions and adult or child face.

<https://doi.org/10.1371/journal.pone.0278483.g001>

the expressions sadness and happiness, there was a higher sensitivity to the expression at an intensity of 25% in adults than child faces (sadness: $Z = -3.477$, $p = .001$; happiness: $Z = 3.1$, $p = .002$), which indicates an own-age bias. For all other expressions and intensities no significant difference in sensitivity to expressions in adult versus child faces was found (all $p > .004$).

In addition, we conducted exploratory analyses on the relation between the age of the participants and their sensitivity or bias. First, we used Kendall's tau (τ) to investigate a non-parametric correlation between age and the mean sensitivity (i.e. medians of sensitivity for all intensities combined; and for both adult and child faces) for the four expressions, tested against alpha 0.0125. There was a positive correlation between age of the participant and sensitivity to a neutral expression ($\tau = .224$, $p < .001$), disgust ($\tau = .196$, $p < .001$), sadness ($\tau = .288$, $p < .001$), but not for happiness ($\tau = .099$, $p = .087$). In addition, we used Kendall's τ to investigate the correlation between age and sensitivity at 25% intensity for the separate emotions and ages of the face, tested against alpha 0.008. There was a positive correlation between age of the participant and sensitivity to disgusted expressions in adult faces ($\tau = .217$; $p < .001$): thus with age, one becomes more sensitive to subtle disgusted expressions displayed by adults. None of the other correlations reached significance (all $p > .01$), although a positive trend was observed for disgusted child faces ($\tau = .106$; $p = .05$). Finally, the Kendall's tau correlation analyses between age and direction of bias (aprime of adult faces minus aprime of child faces) for each of the four expressions at 25% intensity, tested against alpha 0.0167, revealed no significant correlations.

Discussion

The current study investigated the presence of an age bias in labelling facial expressions of emotion in typically developing children and adults. Specific focus was on the influence of different expressions (related to disgust, sadness, happiness) and different intensities (0, 25, 50, 75, 100%) on age bias. The results show that children and adults have a bias at low levels of intensity (i.e. 25%). In the case of sadness, children have an age bias for all intensities. As such, the impact of the age of the face seems largest for expressions which might be most difficult to recognise: expressions displayed at 25% intensity and sadness. Moreover, it appears that both adults and children label expressions best in adult rather than child faces (except for children's rating of disgust). This results in an own-age bias for adults but an other-age bias for children in the labelling of facial expressions.

The current findings expand previous research on age-biases in labeling facial expressions. Although an own-age bias has been shown to be present in adults [15], previous findings in children are conflicting ([18] versus [19]) but research did not investigate this bias in low intense expressions. The current findings reveal that children have an age bias, but that it is mainly present for subtle and sad expressions. As such, the results are partly in line with both studies: it confirms the general conclusion of the existence of an own-age bias by Haushild and colleagues [19], but replicates Vetter and colleagues [18] in that for most expressions this bias is absent for expressions with high intensity. As such, the current findings reveal that expressions with low intensities are not only more difficult to label than higher expressions [9] but labeling these expressions is also more susceptible to the age of the face. Similarly, while it is known that some expressions are more difficult to label than others [21], the present study suggests that particularly the difficult expression of sadness is subject to age bias.

Why would the age of the face affect labelling a facial expression particularly for expressions and intensities that are more difficult to label? In facial expressions that are easy to label (such as happy or high intensive expressions) the facial features are likely more salient: the stimulus is more conspicuous and enhances more sensory gain, and is thus more accessible to the perceptual system as well as capturing more attention [31]). Moreover, the facial features are more distinctive: they are unique to a specific expression [31]. As such, the signal-to-noise ratio can be expected to be high for these expressions. On the contrary, for low-intense expressions the signal-to-noise ratio is very low. Here, any further reduction of signal or increase of noise significantly hampers the ability to label the expression. Children's facial expressions of sadness and happiness are rated to be slightly less clear than adult's expressions, but no difference is reported for disgust (rated by adults: [28]; rates by children only available for child faces: [29]). This slightly reduced clarity could decrease the signal and as such the signal-to-noise ratio, causing a bias for adult expressions of sadness and happiness at low intensities in both adults and children. A component that increases the signal-to-noise ratio is experience [12, 31]: more frequent exposure to a facial expression enhances the ability to process and consequently label the expression. For instance, 'natural' differences in the level of experience with specific expressions due to abusive parents affect the ability to label such expressions [32]. Moreover, increased experience with subtle expressions through training increases sensitivity [33]. Children arguably have less experience with facial expressions than adults. As a consequence, expressions that are difficult to process (i.e. sadness; [21]) might be particularly susceptible to decreases in the signal such as when they are presented on a child's face. Overall, it can be proposed that the signal-to-noise ratio, affected by the face's age, the facial expression and the experience of the participant, at least partly explains why labeling the expression seems most difficult in low-intense expressions on children's faces.

The presence of a bias on labeling of facial expressions can be placed in context of the wider range of components that make up emotion reasoning [12]. These components develop

throughout childhood, but the order in which they are primarily tested (and might emerge) is discrimination, followed by intermodal matching, categorization, event-emotion matching and social referencing, and finally labelling [12]. Moreover, several behavioural experiments testing discrimination or categorization already require a participant to detect, attend to, and remember the facial configuration [12]. Furthermore, the emergence of these components relates to the development of other processes, such as sensory maturation, memory, attention, and knowing emotional words [12]. As such, if there are biases in either the components preceding labeling, or in other processes that play a role in emotion reasoning, these likely affect labeling of expressions as well. Indeed, working memory for emotional faces already has a response bias to happy faces [34]. Moreover, multiple studies have shown that specific emotional faces are detected faster than others when presented amongst neutral faces, although there is a debate on whether this so-called superiority effect is mostly present for happy or angry faces [35–37]. Interestingly, for detection speed there is no own-age bias in children, nor for happy faces in adult participants. However, this bias was observed for angry and fearful faces in adult participants [37]. Although this implies that processes underlying labeling of expressions are already affected by a bias towards specific expressions or stimulus age, it is important to realize that the biases in detection concern processing speed rather than the accuracy that was the focus of the current task, and that working memory likely plays a minimal role in labeling when stimulus and labels are presented at the same time. As such, future research should reveal whether the observed age-biases for labeling of specific emotional expressions are (partly) due to biases in underlying processes.

The current results have implications for social situations including children and adults in which the focus lies on emotion reasoning. In situations such as training, advertisement and movies, where children and adults need to quickly respond to facial expressions, it is important to consider the age biases found in current study. For example, training emotion reasoning of disgusted faces in children would be more effective by starting with pictures of children and subsequently other age-groups. This is currently often not incorporated: most emotion recognition training programs present adult faces, even if aimed at children [38–40].

This study has major strengths. To our current knowledge, it is the first report that investigated an age bias in labeling facial expressions in faces with a range of intensities and expressions, in both children and adults. Moreover, it includes a large group of children that results in high power and allowed exploration of age differences in labelling of these expressions. Furthermore, it applies appropriate statistical tests robust for the observed non-normal distributions in the data. Nevertheless, some limitations need to be kept in mind while interpreting the current results. A possible limitation is that emotion reasoning in daily life is not directly comparable to the lab: in daily life additional information from context, words and body postures or movements aids emotion reasoning [10, 12, 14]. On the other hand, emotion reasoning is hampered in daily life by a wider set of expressions that someone can possibly display, many more than the four expressions participants could choose from in the current study. As such, emotion reasoning in the context of pictures presented in a computer task cannot be fully generalized to emotion reasoning in daily life. Furthermore, we did not control for differences between stimuli in low-level properties, such as spatial frequency, brightness, or contrast. Low-level properties play an important role in the processing and labeling of emotional faces, as sensitivity to several properties continues to develop throughout childhood [41], and such properties are used differently in different age-groups for processing emotional expressions [42–45]. In fact, the correction of distortions resulting from the morphing might have introduced more higher spatial frequencies (represented in edges) and removed lower spatial frequencies (represented in blurry overlap that was corrected). It can thus not be excluded that the observed effects are due to differences in low-level properties between the stimuli instead

of the expression-label itself, nor that the manual stimulus corrections influenced part of these effects. In addition, we observed that a lot of participants in the current study consistently scored very well or very badly, the so-called ceiling and floor effects. Follow-up research should consider using a wider range of intensities, particularly between 0 and 50%, to get a better grasp of sensitivity to subtle facial expressions. Relatedly, the current study presents four trials per condition. Although this is low compared to studies in adults, primary studies on development of labelling expressions with different intensities presented only two trials per condition [9, 17]. Nevertheless, four trials limit the possible variance within each participant, and allows for conclusions on a limited optional outcome in sensitivity. Nevertheless, the current study still yields high power with this number of trials [27]. Another limitation is that the sample size of the child group is much larger than of the adult group. The reasons for this discrepancy are: that this study focused on children and included adults primarily for comparison of conclusions; that the group of children was large to allow studying effects of age; and that the child sample was collected as part of a museum exhibition in which we wanted to allow any child to take part in. Nevertheless, one should note that even the adult sample is large enough to yield high power in the current experimental set-up [27].

In conclusion, both children and adults exhibit an age bias in labeling subtle facial expressions of emotions. It is thus important for studies on emotion reasoning and in practical situations in which one wants the viewer to label a facial expression (such as clinical training, advertisement, or movies) to take the age of the actor into account.

Supporting information

S1 Appendix.
(DOCX)

Author Contributions

Conceptualization: Dafni Surian, Carlijn van den Boomen.

Data curation: Dafni Surian, Carlijn van den Boomen.

Formal analysis: Dafni Surian, Carlijn van den Boomen.

Funding acquisition: Carlijn van den Boomen.

Investigation: Dafni Surian, Carlijn van den Boomen.

Methodology: Dafni Surian, Carlijn van den Boomen.

Project administration: Carlijn van den Boomen.

Resources: Carlijn van den Boomen.

Software: Carlijn van den Boomen.

Supervision: Carlijn van den Boomen.

Validation: Carlijn van den Boomen.

Visualization: Dafni Surian, Carlijn van den Boomen.

Writing – original draft: Dafni Surian.

Writing – review & editing: Dafni Surian, Carlijn van den Boomen.

References

1. Boyatzis C. J., Chazan E., & Ting C. Z. (1993). Preschool children's decoding of facial emotions. *The Journal of genetic psychology*, 154(3), 375–382. <https://doi.org/10.1080/00221325.1993.10532190> PMID: 8245911
2. Junge C., Valkenburg P.M., Dekovic M., & Branje S. (2020). The building blocks of social competence: contributions of the Consortium Individual Development. *Developmental Cognitive Neuroscience*, 45, <https://doi.org/10.1016/j.dcn.2020.100861>
3. Pereira M. R., Barbosa F., de Haan M., & Ferreira-Santos F. (2019). Understanding the development of face and emotion processing under a predictive processing framework. *Developmental Psychology*, 55(9), 1868–1881. <https://doi.org/10.1037/dev0000706> PMID: 31464491
4. Grossmann T., Striano T., & Friederici A. D. (2007). Developmental changes in infants' processing of happy and angry facial expressions: A neurobehavioral study. *Brain and cognition*, 64(1), 30–41. <https://doi.org/10.1016/j.bandc.2006.10.002> PMID: 17169471
5. LaBarbera J. D., Izard C. E., Vietze P., & Parisi S. A. (1976). Four- and six-month-old infants' visual responses to joy, anger, and neutral expressions. *Child Development*, 535–538. <https://doi.org/10.2307/1128816> PMID: 1269322
6. Leppänen J. M., & Nelson C. A. (2009). Tuning the developing brain to social signals of emotions. *Nature Reviews Neuroscience*, 10(1), 37–47. <https://doi.org/10.1038/nrn2554> PMID: 19050711
7. Walker-Andrews A. S. (1997). Infants' perception of expressive behaviors: differentiation of multimodal information. *Psychological bulletin*, 121(3), 437. <https://doi.org/10.1037/0033-2909.121.3.437> PMID: 9136644
8. Durand K., Gallay M., Seigneuric A., Robichon F., & Baudouin J. Y. (2007). The development of facial emotion recognition: The role of configural information. *Journal of experimental child psychology*, 97(1), 14–27. <https://doi.org/10.1016/j.jecp.2006.12.001> PMID: 17291524
9. Gao X., & Maurer D. (2009). Influence of intensity on children's sensitivity to happy, sad, and fearful facial expressions. *Journal of experimental child psychology*, 102(4), 503–521. <https://doi.org/10.1016/j.jecp.2008.11.002> PMID: 19124135
10. Barrett L.F., Adolphs R., Marsella S., Martinez A.M., & Pollak S.D. (2019). Emotional expressions reconsidered: challenges to inferring emotion from human facial movements. *Psychological science*, 20(11), 1–68. <https://doi.org/10.1177/1529100619832930> PMID: 31313636
11. Hoemann K., Wu R., LoBue V., Oakes L.M., Fei X., & Barrett L.F. (2020). Developing an understanding of emotion categories: lessons from objects. *Trends in Cognitive Sciences*, 24(1), 39–51. <https://doi.org/10.1016/j.tics.2019.10.010> PMID: 31787499
12. Ruba A.L., & Pollak S.D. (2020). The development of emotion reasoning in infancy and early childhood. *Annual review in developmental psychology*, 2, 503–531. <https://doi.org/10.1146/annurev-devpsych-060320-102556>
13. Barrett L. F., Mesquita B., & Gendron M. (2011). Context in emotion perception. *Current Directions in Psychological Science*, 20(5), 286–290. <https://doi.org/10.1177/0963721411422522>
14. Coulson M. (2004). Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *Journal of nonverbal behavior*, 28(2), 117–139. <https://doi.org/10.1023/B:JONB.0000023655.25550.be>
15. Riediger M., Voelkle M. C., Ebner N. C., and Lindenberger U. (2011). Beyond "happy, angry, or sad?": age-of-poser and age-of-rater effects on multi-dimensional emotion perception. *Cognitive Emotion*, 25, 968–982. <https://doi.org/10.1080/02699931.2010.540812> PMID: 21432636
16. Rhodes M. G., & Anastasi J. S. (2012). The own-age bias in face recognition: a meta-analytic and theoretical review. *Psychological bulletin*, 138(1), 146. <https://doi.org/10.1037/a0025750> PMID: 22061689
17. Griffiths S., Penton-Voak I. S., Jarrold C., & Munafò M. R. (2015). No Own-Age Advantage in Children's Recognition of Emotion on Prototypical Faces of Different Ages. *PLoS One* 10(5), e0125256. <https://doi.org/10.1371/journal.pone.0125256> PMID: 25978656
18. Vetter N. C., Drauschke M., Thieme J., & Altgassen M. (2018). Adolescent basic facial emotion recognition is not influenced by puberty or own-age bias. *Frontiers in psychology*, 9, 956. <https://doi.org/10.3389/fpsyg.2018.00956> PMID: 29977212
19. Hauschild K. M., Felsman P., Keifer C. M., & Lerner M. D. (2020). Evidence of an Own-Age Bias in Facial Emotion Recognition for Adolescents With and Without Autism Spectrum Disorder. *Frontiers in Psychiatry*, 11, 428. <https://doi.org/10.3389/fpsyg.2020.00428> PMID: 32581859
20. Griffiths S., Jarrold C., Penton-Voak I. S., Woods A. T., Skinner A. L., & Munafò M. R. (2017). Impaired recognition of basic emotions from facial expressions in young people with autism spectrum disorder: Assessing the importance of expression intensity. *Journal of autism and developmental disorders*, 1–11. <https://doi.org/10.1007/s10803-017-3091-7>

21. Herba C., & Phillips M. (2004). Annotation: Development of facial expression recognition from childhood to adolescence: Behavioural and neurological perspectives. *Journal of Child Psychology and Psychiatry*, 45(7), 1185–1198. <https://doi.org/10.1111/j.1469-7610.2004.00316.x> PMID: 15335339
22. Lawrence K., Campbell R., & Skuse D. (2015). Age, gender, and puberty influence the development of facial emotion recognition. *Frontiers in psychology*, 6, 761. <https://doi.org/10.3389/fpsyg.2015.00761> PMID: 26136697
23. Kreider C. M., Bendixen R. M., Young M. E., Prudencio S. M., McCarty C., & Mann W. C. (2016). Social networks and participation with others for youth with learning, attention, and autism spectrum disorders: Réseaux sociaux et participation avec les autres, chez des adolescents ayant des troubles d'apprentissage, de l'attention et du spectre de l'autisme. *Canadian Journal of Occupational Therapy*, 83(1), 14–26. <https://doi.org/10.1177/0008417415583107> PMID: 26755040
24. Shanok N. A., Jones N. A., & Lucas N. N. (2019). The nature of facial emotion recognition impairments in children on the autism spectrum. *Child Psychiatry & Human Development*, 50(4), 661–667. <https://doi.org/10.1007/s10578-019-00870-z>
25. Teunisse J. P., & de Gelder B. (2001). Impaired categorical perception of facial expressions in high-functioning adolescents with autism. *Child Neuropsychology*, 7(1), 1–14. <https://doi.org/10.1076/chin.7.1.1.3150> PMID: 11815876
26. Berggren S., Fletcher-Watson S., Milenkovic N., Marschik P. B., Bölte S., & Jonsson U. (2018). Emotion recognition training in autism spectrum disorder: A systematic review of challenges related to generalizability. *Developmental neurorehabilitation*, 21(3), 141–154. <https://doi.org/10.1080/17518423.2017.1305004> PMID: 28394669
27. Baker D. H., Vilidaitė G., Lygo F. A., Smith A. K., Flack T. R., Gouws A. D., et al. (2021). Power contours: Optimising sample size and precision in experimental psychology and human neuroscience. *Psychological Methods*, 26(3), 295–314. <https://doi.org/10.1037/met0000337> PMID: 32673043
28. Langner O., Dotsch R., Bijlstra G., Wigboldus D. H., Hawk S. T., & Van Knippenberg A. D. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and emotion*, 24(8), 1377–1388. <https://doi.org/10.1080/02699930903485076>
29. Bijsterbosch G., Mobach L., Verpaalen I. A., Bijlstra G., Hudson J. L., Rinck M., et al. (2021). Validation of the child models of the Radboud Faces Database by children. *International Journal of Behavioral Development*, 0165025420935631. <https://doi.org/10.1177/0165025420935631>
30. Craig A. (1979). Nonparametric measures of sensory efficiency for sustained monitoring tasks. *Human Factors*, 21(1), 69–77. <https://doi.org/10.1177/001872087902100109> PMID: 468268
31. Calvo M.G. & Nummenmaa L. (2016). Perceptual and affective mechanisms in facial expression recognition: An integrative review. *Cognition and Emotion*, 30(6) 1081–1106, <https://doi.org/10.1080/02699931.2015.1049124> PMID: 26212348
32. Pollak S. D., & Sinha P. (2002). Effects of early experience on children's recognition of facial displays of emotion. *Developmental psychology*, 38(5), 784. <https://doi.org/10.1037/0012-1649.38.5.784> PMID: 12220055
33. Pollux P. M. J. (2016). Improved categorization of subtle facial expressions modulates late positive potential. *Neuroscience*, 322, 152–163. <https://doi.org/10.1016/j.neuroscience.2016.02.027> PMID: 26912280
34. Tamm G., Kreegipuu K., Harro J., & Cowan N. (2017). Updating schematic emotional facial expressions in working memory: Response bias and sensitivity. *Acta psychologica*, 172, 10–18. <https://doi.org/10.1016/j.actpsy.2016.11.002> PMID: 27835749
35. Hodson S., Viding E. & Lavie N. (2011). Attentional capture by irrelevant emotional distractor faces. *Emotion*, 11(2), 346. <https://doi.org/10.1037/a0022771> PMID: 21500903
36. Lundqvist D. & Ohman A. Emotion regulates attention: The relation between facial configurations, facial emotion, and visual attention. *Vis. Cogn.* 12(1), 51–84 (2005).
37. Zsido AN, Arato N, Ihasz V, Basler J, Matuz-Budai T, Inhof O, et al. (2021). "Finding an Emotional Face" Revisited: Differences in Own-Age Bias and the Happiness Superiority Effect in Children and Young Adults. *Front. Psychol.* 12:580565. <https://doi.org/10.3389/fpsyg.2021.580565> PMID: 33854456
38. Golan O., Granader E., McClintock S., Day K., Leggett V., & Baron—Cohen S. (2010). Enhancing emotion recognition in children with autism spectrum condition: an intervention using animated vehicles with real emotional faces. *Journal of Autism and Developmental Disorders*, 40(3), 269–279. <https://doi.org/10.1007/s10803-009-0862-9> PMID: 19763807
39. Hopkins I. M., Gower M. W., Perez T.A., Smith D. S., Amthor F. R., Wimsatt F. C., et al. (2011). Avatar assistant: improving social skills in students with an ASD through a computer-based intervention. *Journal of Autism and Developmental Disorders*, 41(11), 1543–1555. <https://doi.org/10.1007/s10803-011-1179-z> PMID: 21287255

40. Ryan C., & Charragáin C. N. (2010). Teaching emotion recognition skills to children with autism. *Journal of Autism and Developmental Disorders*, 40(12), 1505–1511. <https://doi.org/10.1007/s10803-010-1009-8> PMID: 20386975
41. van den Boomen C., van der Smagt M.J., & Kemner C. (2012). Keep your eyes on development: the behavioral and neurophysiological development of visual mechanisms underlying form processing. *Frontiers in Psychiatry*, 3, <https://doi.org/10.3389/fpsy.2012.00016> PMID: 22416236
42. Jessen S., & Grossmann T. (2017). Exploring the role of spatial frequency information during neural emotion processing in human infants. *Frontiers in Human Neuroscience*, 11. <https://doi.org/10.3389/fnhum.2017.00486> PMID: 29062275
43. Peters J.C., & Kemner C. (2017). Facial expressions perceived by the adolescent brain: Towards the proficient use of low spatial frequency information. *Biological Psychology*, 129, 1–7. <https://doi.org/10.1016/j.biopsycho.2017.07.022> PMID: 28778549
44. van den Boomen C., Munsters N. M., & Kemner C. (2019). Emotion processing in the infant brain: The importance of local information. *Neuropsychologia*, 126, 62–68. <https://doi.org/10.1016/j.neuropsychologia.2017.09.006> PMID: 28889996
45. Vlamings P. H. J. M., Jonkman L. M., van Daalen E., van der Gaag R. J., & Kemner C. (2010). Basic abnormalities in visual processing affect face processing at an early age in autism spectrum disorder. *Biological Psychiatry*, 68(12), 1107–1113. <https://doi.org/10.1016/j.biopsych.2010.06.024> PMID: 20728876