

Tracking the recognition of static and dynamic facial expressions of emotion across the life span

Anne-Raphaëlle Richoz

Eye and Brain Mapping Laboratory (iBMLab),
Department of Psychology, University of Fribourg,
Fribourg, Switzerland

LPNC, University of Grenoble Alpes, Grenoble, France



Junpeng Lao

Eye and Brain Mapping Laboratory (iBMLab),
Department of Psychology, University of Fribourg,
Fribourg, Switzerland

Olivier Pascalis

LPNC, University of Grenoble Alpes, Grenoble, France

Roberto Caldara

Eye and Brain Mapping Laboratory (iBMLab),
Department of Psychology, University of Fribourg,
Fribourg, Switzerland



The effective transmission and decoding of dynamic facial expressions of emotion is omnipresent and critical for adapted social interactions in everyday life. Thus, common intuition would suggest an advantage for dynamic facial expression recognition (FER) over the static snapshots routinely used in most experiments. However, although many studies reported an advantage in the recognition of dynamic over static expressions in clinical populations, results obtained from healthy participants are contrasted. To clarify this issue, we conducted a large cross-sectional study to investigate FER across the life span in order to determine if age is a critical factor to account for such discrepancies. More than 400 observers (age range 5–96) performed recognition tasks of the six basic expressions in static, dynamic, and shuffled (temporally randomized frames) conditions, normalized for the amount of energy sampled over time. We applied a Bayesian hierarchical step-linear model to capture the nonlinear relationship between age and FER for the different viewing conditions. Although replicating the typical accuracy profiles of FER, we determined the age at which peak efficiency was reached for each expression and found greater accuracy for most dynamic expressions across the life span. This advantage in the elderly population was driven by a significant decrease in performance for static images, which was twice as large as for the young adults. Our data posit the use of dynamic stimuli as being critical in the assessment of FER in the elderly population, inviting caution when drawing

conclusions from the sole use of static face images to this aim.

Introduction

Human faces convey a wealth of dynamic signals that are critical for an adequate and rapid categorization of the emotional states of others. Yet the vast majority of studies investigating expression recognition have relied on static images that commonly display the apex or the highest state of a given expression. In everyday life, however, facial expressions are rarely transmitted and decoded through static snapshots of internal states. Natural human interactions are a highly dynamic (and multimodal) phenomenon with faces evolving over time while transmitting distinct signals to convey diverse emotional states. Dynamic expressions provide observers with additional cues related to their inherent temporal properties, such as their unfolding speed (slow vs. fast; Bould & Morris, 2008; Bould, Morris, & Wink, 2008; Kamachi et al., 2001), rise time (from the neutral to the highest state; R. E. Jack, Garrod, & Schyns, 2014; Recio, Schacht, & Sommer, 2013) or intensity (Bould et al., 2008), critical for an adequate categorization. Therefore, dynamic faces are richer and ecologically more valid depictions of the way expressions are encountered in everyday life compared to static images (e.g., Johnston, Mayes, Hughes, &

Citation: 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(9):5, 1–27, <https://doi.org/10.1167/18.9.5>.

<https://doi.org/10.1167/18.9.5>

Received February 27, 2018; published September 7, 2018

ISSN 1534-7362 Copyright 2018 The Authors



Young, 2013; Paulmann, Jessen, & Kotz, 2009; Trautmann, Fehr, & Herrmann, 2009). Interestingly, from an evolutionary perspective, humans have had more experience with dynamic faces as static pictures only appeared during the last century with the advent of photography and the rapid expansion of digital tools and social networks. The decoding of static faces is also a learnt behavior that develops throughout life. As the human visual system is, from birth on, steadily stimulated by dynamic signals from faces with minimal exposure to static faces, common intuition would suggest the existence of a particular expertise to decode such events with the presence of an advantage for the recognition of dynamic over static expressions.

Previous studies that have attempted to investigate this question have yielded inconsistent findings (for a review, see Alves, 2013; Fiorentini & Viviani, 2011; Kätsyri, 2006; Krumhuber, Kappas, & Manstead, 2013). Some behavioral studies have revealed an advantage (e.g., Ambadar, Schooler, & Cohn, 2005; Cunningham & Wallraven, 2009; Giard & Peronnet, 1999; Knappmeyer, Thornton, & Bülthoff, 2003; Paulmann et al., 2009; Wehrle, Kaiser, Schmidt, & Scherer, 2000), whereas others have revealed that the benefits of dynamic cues in facial expression recognition may be minimal (e.g., Gold et al., 2013) or inexistent (e.g., Fiorentini & Viviani, 2011). These contrasting findings suggest that the dynamic advantage for facial expression recognition is not as straightforward as it may appear. Rather, it seems that the physical properties of the stimuli presented as well as clinical or neuropsychological conditions influence the extent to which dynamic displays lead to processing benefits (Ambadar et al., 2005; Bould et al., 2008; Wallraven, Breidt, Cunningham, & Bülthoff, 2008).

Several studies have shown that the beneficial effects of dynamic events are particularly relevant in suboptimal situations in which the physical information available is limited (Ambadar et al., 2005; Bould et al., 2008), deteriorated, or blurred (Ehrlich, Schiano, & Sheridan, 2000; Kätsyri & Sams, 2008; Wallraven et al., 2008). For example, Wallraven et al. (2008) found that dynamic events increased recognition accuracy of computer-animated facial expressions whose texture or shape were systematically degraded. Similarly, by comparing the ability of observers to recognize expressions from schematic and natural faces, Kätsyri and Sams (2008) and Ehrlich et al. (2000) discovered a recognition advantage for dynamic expressions with schematic but not natural faces. Along the same lines, other studies have revealed that dynamic events provide compensatory cues when subtle facial expressions are presented (Ambadar et al., 2005; Bould et al., 2008). With subtle expressions, additional temporal information may be essential to disambiguate the uncertainty introduced by the lack of intensity.

Similarly, an advantage is noticeable when it comes to clinical conditions as dynamic information provides compensatory cues in suboptimal situations. Dynamic presentations facilitate the recognition of facial expressions in adults and children with intellectual disability (Harwood, Hall, & Shinkfield, 1999), pervasive developmental disorder (Uono, Sato, & Toichi, 2010), and autism (Back, Ropar, & Mitchell, 2007; Gepner, Deruelle, & Grynfeldt, 2001; Tardif, Lainé, Rodriguez, & Gepner, 2007; but see Kätsyri, Saalasti, Tiippana, von Wendt, & Sams, 2008, for Asperger syndrome). In neuropsychology, several brain-injury studies have shown increased recognition performance when dynamic expressions were used (Adolphs, Tranel, & Damasio, 2003; Humphreys, Donnelly, & Riddoch, 1993; Richoz, Jack, Garrod, Schyns, & Caldara, 2015). For example, Humphreys et al. (1993) reported the case of an agnostic patient who was significantly impaired at identifying facial identity and facial expressions when exposed to static images. In contrast, his performance was proficient when asked to judge a subset of facial expressions (i.e., smiling, frowning, or surprise) from dynamic faces animated by light dots. On the same line, we recently investigated the ability of a prosopagnosic patient—the well-studied case of PS—with multiple and extensive brain lesions in the occipitotemporal cortex to recognize facial expressions from static and dynamic faces. Our findings revealed that the patient PS was selectively impaired in decoding static expressions while showing normal performance for the decoding of dynamic emotional expressions. This observation favors the existence of distinct representational systems for static and dynamic expressions or dissociable cortical pathways to access them (Richo et al., 2015). Noteworthy, the advantage for processing dynamic faces in PS is related to a suboptimal information use for static (i.e., bias toward the mouth) compared to dynamic faces (i.e., all face features; Fiset et al., 2017).

Although several neuropsychological studies have shown that the dynamic properties of human facial expressions provide significant processing advantages, other behavioral studies involving healthy observers suggest that this might not be the case (Bould & Morris, 2008, for expressions of high intensity; Christie & Bruce, 1998; Fiorentini & Viviani, 2011; Gold et al., 2013; Jiang et al., 2014; Kamachi et al., 2001, experiment 2). By using a threshold model, Fiorentini and Viviani (2011), for example, reported that neither reaction times nor identification accuracy were more accurate for the dynamic as compared to the static expressions. Similar findings were reported in a later study by Gold et al. (2013). Their results revealed that recognition rates were nearly identical when participants were exposed to static, dynamic, shuffled (temporally randomized expressions), or reversed

expressions. This suggests that the temporal properties provided by moving faces are not necessary for observers to reliably categorize emotional expressions. Altogether, these studies suggest that a healthy visual system seems to be powerful enough to efficiently recognize intense expressions from static faces, leaving only a nonsignificant benefit to the processing of dynamic facial expressions. By contrast, in clinical conditions, the muscular movements associated with the temporal unfolding of an expression may force the observers to shift their attention to different facial features. This may enhance attention and motor simulations (A. Wood, Lupyan, Sherrin, & Niedenthal, 2016; A. Wood, Rychlowska, Korb, & Niedenthal, 2016) in fragile or neurologically impaired face systems, which may explain the increased performance with dynamic signals in these populations.

Interestingly, there are stages in healthy observers during which the perceptual system is also particularly fragile or immature. For example, from early infancy to late adolescence, the brain undergoes a wide array of anatomical and functional changes as it develops (e.g., Blakemore, 2012; Blakemore & Choudhury, 2006; Casey, Tottenham, Liston, & Durston, 2005; Durston et al., 2001). Similarly, during normal aging, the cognitive functions decline, which is induced by age-related loss of synaptic contacts, neural apoptosis (e.g., Raz, 2000; Rossini, Rossi, Babiloni, & Polich, 2007), reduction in cerebral blood flow (e.g., Chen, Rosas, & Salat, 2011), or volume reduction in different brain regions (e.g., amygdala, hippocampus, frontal cortex; Calder et al., 2003; C. R. Jack et al., 1997; Navarro & Gonzalo, 1991; Ruffman, Henry, Livingstone, & Phillips, 2008). Considering the increased vulnerability of the brain under neural architectural changes (Andersen, 2003; Hof & Morrison, 2004), it is possible that healthy young children and normal aging adults also benefit from the presentation of dynamic faces. However, only a few developmental studies have compared facial expression recognition in children using both static and dynamic stimuli (Nelson, Hudspeth, & Russell, 2013; Nelson & Russell, 2011). These studies yielded equivocal results, none of them revealing a significant advantage for dynamic over static stimuli; two studies even pointed to differences favoring static stimuli (Nelson & Russell, 2011; Widen & Russell, 2015). Nevertheless, most of these studies tested facial expression recognition with the use of a single actor and provided additional information about face, body movements, and vocal intonations, which may have facilitated expression recognition. In the aging literature, a small number of studies examined facial expression recognition with static and dynamic faces (Grainger, Henry, Phillips, Vanman, & Allen, 2015; Krendl & Ambady, 2010; Sze, Goodkind, Gyurak, & Levenson, 2012). Although most of these

studies pointed to a dynamic advantage for the recognition of facial expressions, they (a) did not use a database of static and dynamic stimuli controlled for the amount of *low-level* visual information carried over time (Grainger et al., 2015; Sze et al., 2012), (b) were limited to a subset of emotional expressions (Krendl & Ambady, 2010), (c) included participants in only one condition (Krendl & Ambady, 2010), or (d) relied on dynamic movies that were not displaying natural expressions (Grainger et al., 2015). These methodological issues considerably limit firm conclusions on the potential benefits of dynamic cues for the recognition of facial expressions in elderly people.

Developmental studies have reported an early tuning to culturally specific expressions (Geangu et al., 2016; for a review, see Caldara, 2017) and emotion-dependent differences in the development of facial expression recognition abilities with some expressions being recognized earlier (e.g., happiness) than others (e.g., fear) (Durand, Gally, Seigneureic, Robichon, & Baudouin, 2007; Gao & Maurer, 2010; Gross & Ballif, 1991; Herba & Phillips, 2004; Rodger, Vizioli, Ouyang, & Caldara, 2015). Similarly, studies with elderly people have shown that the recognition of some expressions decreases with increasing age, and the recognition of others remains stable or even improves (Calder et al., 2003; MacPherson, Phillips, & Della Sala, 2002; Sullivan & Ruffman, 2004b; Zhao, Zimmer, Shen, Chen, & Fu, 2016). Most of these studies were, however, conducted with static posed images, and only little is known about the effects of aging on the recognition of genuine dynamic emotional expressions.

To fill this gap in the developmental literature, we investigated whether the advantage for dynamic stimuli extends to other populations with immature (i.e., young children) or fragile (i.e., elderly adults) face-processing systems. We conducted a large cross-sectional study involving more than 400 observers (age range 5–96) in order to investigate facial expression recognition from early to elderly age. Observers performed categorization tasks of the six basic expressions (anger, disgust, fear, happiness, sadness, and surprise) in three conditions: static, dynamic, and shuffled (temporally randomized frames; Gold et al., 2013). Importantly, we relied on a specific database of static, dynamic, and shuffled stimuli created by Gold et al. (2013). Our experimental choice was driven by the fact that these authors also used an ideal observer model to objectively measure the amount of *low-level* physical information carried by the stimuli. It is worth noting that most studies investigating the presence of a dynamic advantage (e.g., Ambadar et al., 2005; Bould & Morris, 2008; Bould et al., 2008; Cunningham & Wallraven, 2009; Fiorentini & Viviani, 2011; Kätsyri & Sams, 2008) directly compared participants' recognition rates in the static and dynamic conditions without control-

ling the amount of low-level information physically available to the observers. As mentioned by Gold et al., the absence of an objective measure of stimulus information makes it difficult, in most cases, to determine whether increased recognition rates are due to adequate categorization skills, to the amount of physical information available, or to a combination of both factors. By comparing human expression recognition scores with the performance of a statistically ideal observer, Gold et al. reported that their dynamic stimuli did not provide additional low-level information than what was already offered by their static snapshots (for additional information, see Gold et al., 2013). In addition to this approach, we modeled the relationship between age and facial-expression recognition by using a hierarchical Bayesian approach with a step-linear model. Our results revealed emotion-specific advantages for dynamic stimuli. More specifically, although participants displayed nearly identical categorization performance for the static and dynamic expressions of fear and sadness, *all* the other expressions were more readily labeled as correct when featuring dynamic displays. Overall, the results of this study provide a comprehensive and detailed view of the way in which static and dynamic expressions are recognized across the human life span.

Material and methods

The experiment script, raw data, and analysis codes are open to access on Github (https://github.com/iBMLab/Static_dynamic).

Participants

A total of 444 healthy observers participated in the current study. Subjects who did not respond at least once to all expressions on the first condition/block were excluded from the analyses ($N = 32$), leaving a total number of 412 participants. Their exclusion is based on the difficulty to determine whether they actually did not recognize the expression presented or did not correctly understand the task. A future research paper will investigate the systematic errors of the participants that were excluded.

We intended to collect data from 20 participants in each age group ranging from 5 to 96 years of age. The groups were comprised as follows: 5- to 6-year-olds ($N = 27$, 17 females), 7- to 8-year-olds ($N = 24$, 17 females), 9- to 10-year-olds ($N = 22$, 11 females), 11- to 12-year-olds ($N = 22$, 14 females), 13- to 14-year-olds ($N = 24$, 10 females), 15- to 16-year-olds ($N = 21$, eight females), 17- to 18-year-olds ($N = 21$, 16 females), 19- to 20-year-

olds ($N = 31$, 27 females). From the age of 21 to the age of 96, six different groups were created: 21- to 30-year-olds ($N = 31$, 23 females), 31- to 40-year-olds ($N = 23$, 13 females), 41- to 50-year-olds ($N = 33$, 22 females), 51- to 60-year-olds ($N = 30$, 18 females), 61- to 80-year-olds ($N = 31$, 25 females), and 81- to 96-year-olds ($N = 40$, 30 females).

All participants had normal or corrected-to-normal vision with no neurological or psychiatric history. Children were recruited from primary and high schools in the area of Fribourg, Switzerland. Parental consent was required for all children under the age of 16. Participants older than 16 were recruited at the University of Fribourg through social networks or advertisements. Observers from the university obtained course credits for their participation. All participants signed a consent form that described the main goals of our experiment.

Elderly people were recruited and tested in senior housing in the Fribourg region. We used the Mini-Mental State Examination (Folstein, 1975) in order to determine the eligibility of the elderly people aged 60 and over. This brief cognitive screening test, which has been extensively used and validated since its creation in 1975, allows the assessment of different cognitive functions, such as memory, orientation, attention, language, and recall, through 11 questions with a maximum score of 30. Elderly people with a score below 24 were excluded from our study ($N = 3$) as this score has been set as the most commonly used cutoff score for cognitive impairment (Mitchell, 2009). The ethical committee of the department of psychology of the University of Fribourg approved the study reported here.

Stimuli

We used the same stimuli as those used by Gold et al. (2013). In order to create their database, Gold et al. asked eight actors (four females) to reproduce the six basic facial expressions of emotion as naturally as possible (i.e., anger, disgust, fear, happiness, sadness, and surprise; Ekman & Friesen, 1976). The facial expressions of emotion started from a neutral state and naturally evolved into a full expression. Their apex state (i.e., the point at which an expression reached its fully articulated state) was determined by two raters.

The dynamic faces evolved from a neutral state to a full-blown expression at a frame rate of 30 frames/s. All expressions reached their apex within 30 frames. If the fully articulated expression was reached before 30 frames, one to four supplementary apex frames were appended, but as the actors were asked to maintain the apex for several seconds, this happened for only seven out of 48 movies (for more details, see Gold et al.,

2013). Faces were presented in black and white and cropped at the hairline to present only the internal facial features. Previous experiments have shown that external features attract children's attention (Leitzke & Pollak, 2016). Moreover, the faces were centered and seen through an oval aperture, which was placed in the middle of a gray-colored background. The borders of the oval aperture were slightly blurred in order to produce a progressive transition between the background and the faces (Gold et al., 2013). The faces were resized from the original experiment and each measured 768 pixels in height and 768 pixels in width. They subtended a visual angle of 12° on the screen at a viewing distance of 65 cm. All faces were equated for luminance and contrast.

Based on these dynamic sequences, Gold et al. (2013) generated two other sets of stimuli: a set of frozen images (static condition) and a set of temporally randomized dynamic frames (shuffled condition) (see Figure 1; supplementary videos related to this article can be found under the specific links). In the static condition, movies were created by taking the apex frame of each dynamic sequence and replicating it 30 times in a row. In the shuffled condition, movies were generated by randomly selecting the individual frames of the dynamic sequences. This condition was originally designed to assess whether human observers were sensitive to the temporal development of an expression over time (i.e., order of frames). The results reported by Gold et al. revealed that recognition efficiency did not significantly differ between the dynamic and shuffled expressions in young adults, suggesting that young adults are insensitive to the temporal properties associated with the unfolding of an expression.

We normalized all stimuli for their low-level properties and the amount of energy sampled over time even for the static condition for every frame. More concretely, the video stimuli were normalized across all frames and all expressions using the SHINE toolbox with the default option (Willenbockel et al., 2010). In order to partly account for the differences in visual input between static and dynamic stimuli, we computed the raw pixel intensity differences between each frame of the dynamic movies. We, thus, added these intensity differences to each frame at random permuted locations in the static images. It is worth noting that normalization after adding the noise would defeat this purpose as the difference between each frame will not be straightforwardly comparable with the natural low-level differences in the dynamic stimuli. However, with our approach, we could ensure that all the frames for all the faces in all conditions have equal low-level properties (luminance and contrast). The stimuli were displayed on a color liquid-crystal display with a resolution of 1,440 × 900 pixels and a refresh rate of 60 Hz. The whole experiment was programmed in

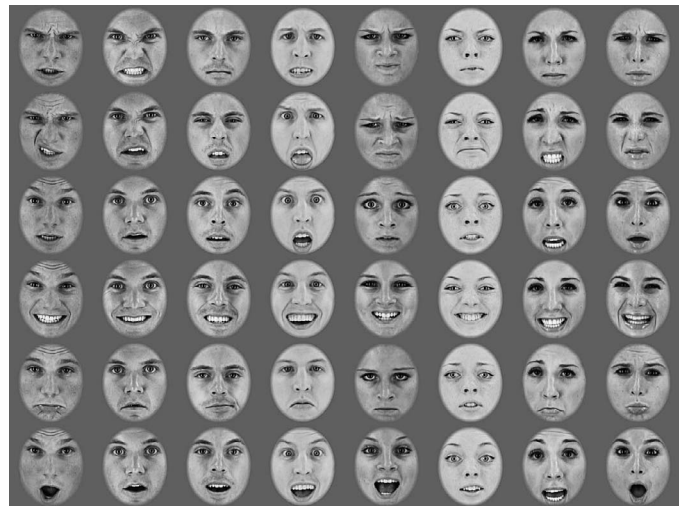


Figure 1. Examples of the stimuli used in our study. Face identities and facial expressions used for the study with each actor (column) and the six expressions (row: anger, disgust, fear, happiness, sadness, surprise). Please note that we inserted noise in the static condition in order to normalize the amount of energy sampled over time across conditions. For an illustrative purpose, see an example for anger for the static (<http://perso.unifr.ch/roberto.caldara/JoV/Anger-static.mov>), dynamic (<http://perso.unifr.ch/roberto.caldara/JoV/Anger-dynamic.mov>), and shuffled conditions (<http://perso.unifr.ch/roberto.caldara/JoV/Anger-shuffled.mov>). The stimuli were adapted with permission from Gold et al. (2013).

MATLAB (MATLAB 2014B; MathWorks, Natick, MA) using the Psychophysics Toolbox (PTB-3; Brainard, 1997; Kleiner et al., 2007).

Procedure

Participants were told that they would see faces expressing different kinds of emotions on a computer screen and their task would be to categorize them as accurately as possible, according to the six following possibilities: anger, disgust, fear, happiness, sadness, and surprise.

In order to familiarize children with the faces and ensure that they understood the conceptual meaning of all expressions, we presented them with printed sheets of the different expressions and asked them to tell us how the person presented on the image was feeling.

All participants sat 65 cm away from a computer screen in a quiet room. Each trial started with a white fixation cross presented at the center of the screen for 500 ms. The stimuli were then presented in a random order, one at a time, in the center of the computer screen for a duration of 1 s each (for a schematic representation of the procedure, see Figure 2). We used the same stimuli presentation time in all three

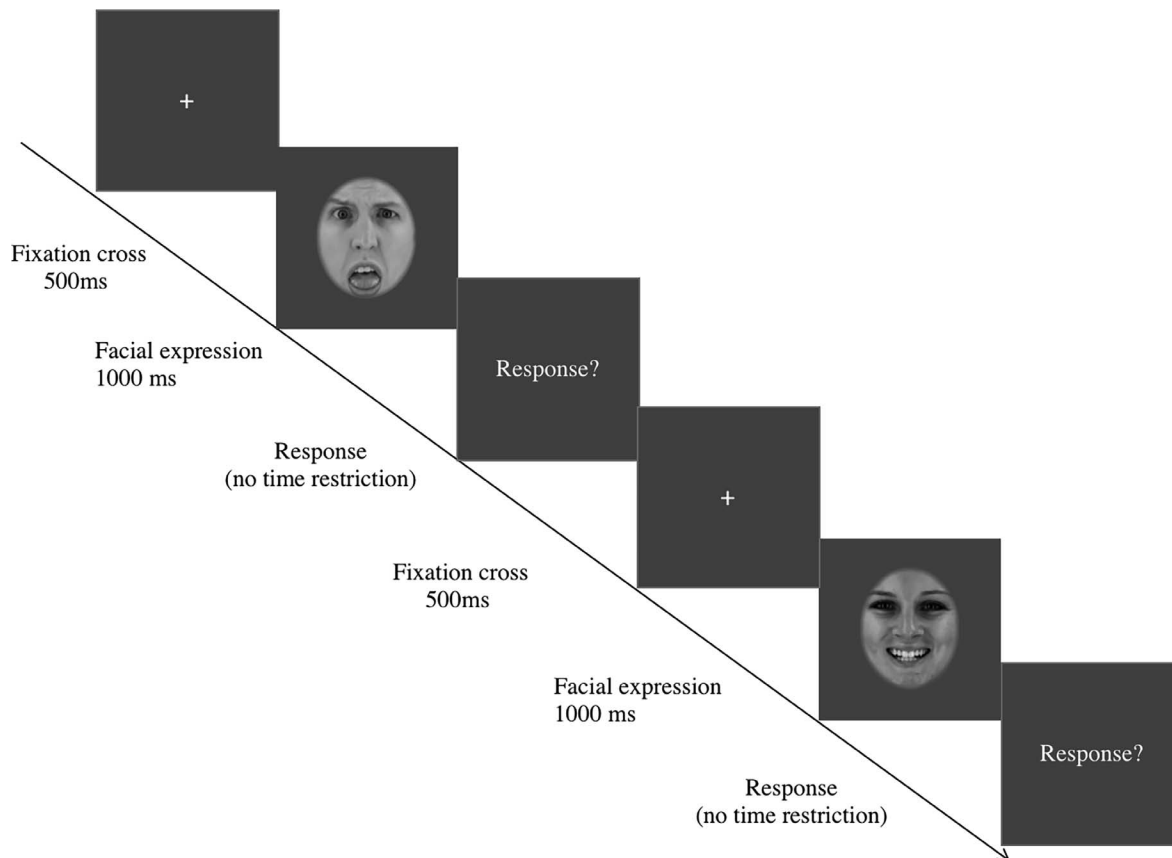


Figure 2. Schematic representation of the procedure. Each trial began with a white fixation cross that was presented for 500 ms, followed by a face presented for 1 s, which expressed one of the six basic facial expressions of emotion: anger, disgust, fear, happiness, sadness, and surprise. After each trial, participants were asked to categorize the previously seen expression.

conditions in order to fully replicate the study by Gold et al. (2013). Note that a presentation time of 1 s was also previously used in other studies with dynamic faces (Adolphs et al., 2003; Recio et al., 2013; Richoz et al., 2015). After each presentation, a response window was displayed on the screen and remained there until the participant answered. Observers categorized each stimulus by using a computer keyboard on which we labeled the keys accordingly. They could press a key labeled “I don’t know” if they were unsure, had not had enough time to see the expression, or did not know the answer. We decided to introduce an “I don’t know” option in order to reduce the noise and response bias produced by the lack of such a key. We gave our participants as much time as required to categorize the expressions and told them that judgment accuracy was important, not the response time. Children under the age of 10, participants who were not familiar with computers, and elderly people over 65 gave their answers verbally to the experimenter who keyed them in. No feedback was provided. The stimuli were blocked by condition. Each condition consisted of two blocks of 48 trials (eight actors, six expressions)

presented twice (96 expressions for each condition) for a total of 288 trials. Participants took part in all three conditions in a counterbalanced random order. The testing was done in one session for adolescents and adults, two or three sessions for participants under 10 or over 65. Before starting the testing phase, participants completed 12 practice trials for each condition.

Data analysis

Data analysis was performed in Python using Jupyter Notebook. Summary statistics by groups are displayed as confusion matrices (Supplementary Figure S1A through E) and line plots (Figure 4) for each condition.

Bayesian modeling was performed using PyMC3 version 3.2, and the results were displayed using Seaborn and Matplotlib. The main aim of the current study was to determine the underlying function between expression-recognition ability and age, conditioned on the basis of different types of visual stimuli. More specifically, we were interested in modeling expression-

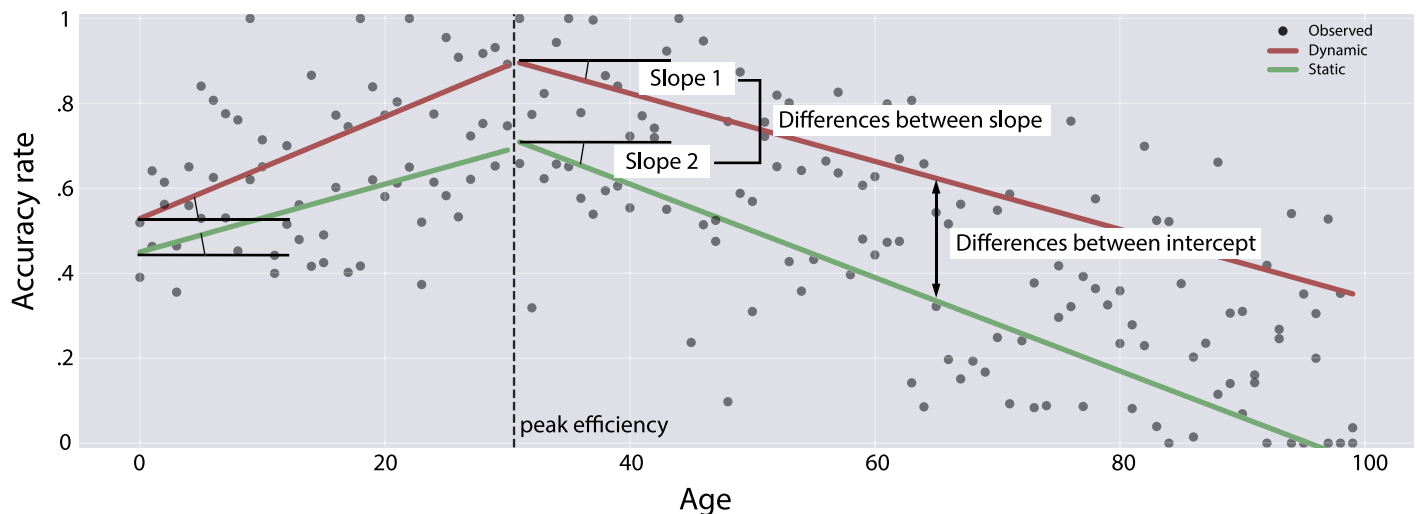


Figure 3. A conceptual representation of the step-linear model. We are interested in the posterior distribution of the peak efficiency and the contrasts between the posterior distribution of different slopes and the different intercepts.

recognition ability as a function of age, expression, and stimuli type (static, dynamic, or shuffled):

$$Recognition\ Ability = f(age, expression, stimuli\ type).$$

Here, recognition abilities were measured using the correct identification (i.e., hit). Importantly, as the

target function f is a nonlinear function, in order to capture the increase and then the decrease in the recognition abilities displayed in the data, we constructed a simple step-linear function with two linear equations. The first equation captures the increase in recognition abilities before a break point, defined as the

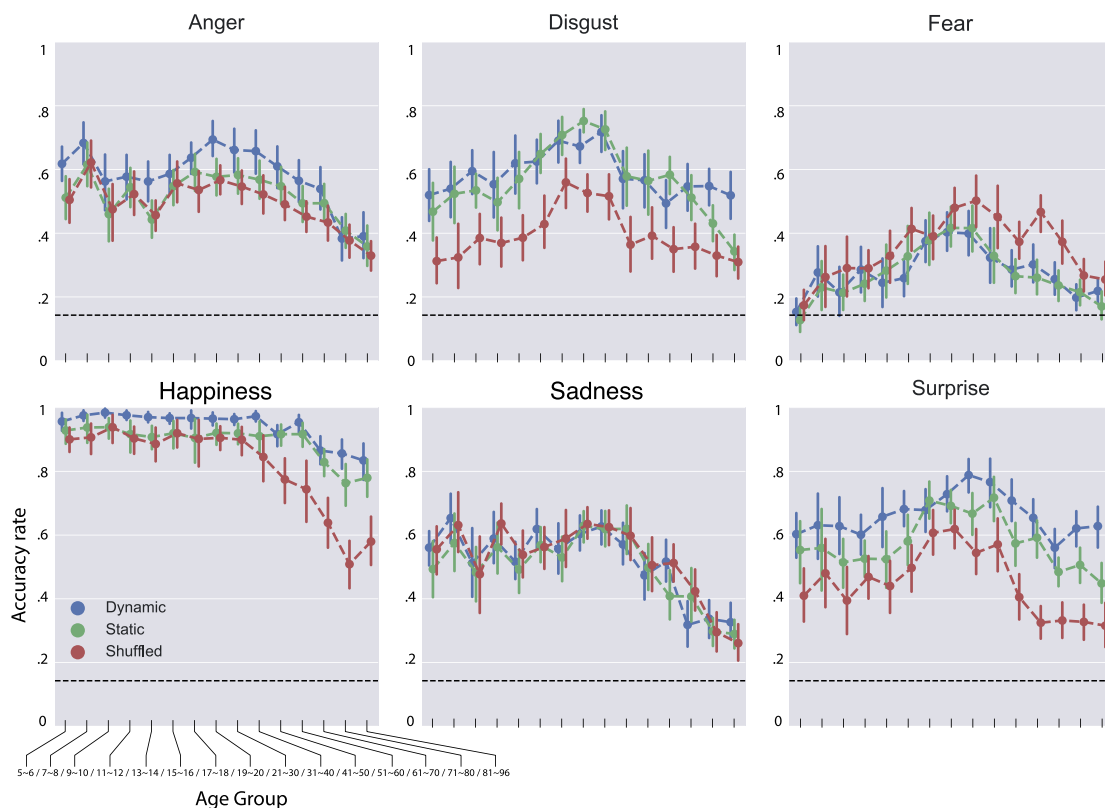


Figure 4. Accuracy across age groups for each expression in the three different conditions. Error bars show 95% bootstrap confidence interval for the mean. Age groups were created as follows: 5–6, 7–8, 9–10, 11–12, 13–14, 15–16, 17–18, 19–20, 21–30, 31–40, 41–50, 51–60, 61–70, 71–80, 81–96.

momentum in age capturing peak efficiency, whereas the second equation captures the decrease in recognition abilities.

$$f_1(\text{age}, \text{expression}, \text{stimuli type}), \text{age} < \tau,$$

$$f_2(\text{age}, \text{expression}, \text{stimuli type}), \text{age} \geq \tau.$$

Here, the break point τ is expressed as a latent variable that is estimated from the model. Both f_1 and f_2 are linear functions of age during the recognition of specific expressions and stimuli type. Thus, the slope of the function f_1 and f_2 (coefficient for age) captures the change in recognition abilities, whereas the intercept of the function captures the general recognition abilities before and after the age τ . We estimated the general dynamic advantage by computing contrasts of the intercepts between the different stimuli type (i.e., static, dynamic, shuffled) and quantified the interaction between stimuli type and age (i.e., whether there is a stronger dynamic advantage in young/old age) by computing contrasts of the slopes. Importantly, considering that the break point τ could occur at different age stages among the expressions, we modeled each expression independently, thus turning the target function into

$$f_1(\text{age}, \text{stimuli type}|\text{expression}), \text{age} < \tau,$$

$$f_2(\text{age}, \text{stimuli type}|\text{expression}), \text{age} \geq \tau,$$

where recognition ability of different types of stimuli is a step-linear function of age conditioned on a specific expression.

In practice, we formulated functions f_1 and f_2 as logistic regressions with the function output being the success probability p in each trial in the binomial distribution. The total number of correct responses for one participant during the presentation of one expression and one stimuli type follows a binomial distribution:

$$k \sim \text{Binomial}(p, n).$$

Thus, this is an extended beta-binomial model with latent variables. The full model is formulated as below:

i for each task (dynamic, static, shuffle), j for each participant.

Hyper-priors of the slope

$$\mu_\beta \sim \text{Student } t(3, 0, 10)$$

$$\sigma_\beta \sim \text{Half Normal}(10)$$

Hyper-prior of the breakpoint

$$\tau \sim \text{Uniform}(0, 100)$$

Hyper-prior of the recognition ability at τ

$$\theta \sim \text{Uniform}(0, 1)$$

$$\kappa \sim \text{Uniform}(0, N\tau)$$

(reparameterized as the mode of beta distribution)

$$a = \theta \times (\kappa - 2) + 1$$

$$b = (1 - \theta) \times (\kappa - 2) + 1$$

For each stimuli type $i \in$ static, dynamic, shuffled:

Prior of the recognition ability at τ

$$\theta_i \sim \text{Beta}(a, b)$$

Priors of the break point

$$\tau_i \sim \text{Normal}(\tau, 10)$$

Priors of the slopes (a indicates before age τ)

$$\beta_i^{a,b} \sim \text{Normal}(\mu_\beta, \sigma_\beta)$$

The intercepts before and after age τ

$$b_i^{a,b} = \text{logit}(\theta_i) - \tau_i \times \beta_i^{a,b}$$

Linear function and invlogit transform

$$\beta_i, \text{Intercept}_i = \begin{cases} \beta_i^a, b_i^a & \text{if } \text{age} < \tau_i \\ \beta_i^b, b_i^b & \text{if } \text{age} \geq \tau_i \end{cases}$$

$$\theta_{i,j} = \beta_i \otimes \text{age}_j + \text{Intercept}_i$$

$$\hat{y}_{i,j} = \text{invlogit}(\theta_{i,j})$$

Observed accurate categorizations

$$k_{i,j} \sim \text{Binomial}(\hat{y}_{i,j}, n_{i,j})$$

As shown above, the slope of each condition is regularized using a weakly informative hyper-prior. The prior of each slope is a normal distribution with the mean distributed as a zero mean Student t distribution with three degrees of freedom and 10 standard deviations and the standard deviations distributed as a half-normal distribution. The hyper-prior of the break point τ is a uniform distribution from 0 to 100, which is the overall mean of the condition-specific break point that follows a normal distribution with 10 standard deviations as prior. Importantly, the intercept of the two linear functions f_1 and f_2 is determined by the recognition ability θ at the break point τ . The condition-specific recognition ability θ_i follows a Beta distribution as prior. Moreover, we reparameterized the beta distribution by the mode θ and the concentration κ (Kruschke, 2014, cf. equation 9.4, p. 223). Here, the mode θ follows a uniform prior between zero and one, and κ follows a uniform prior with two as minimum and the number of trials as maximum.

The probabilistic model was built using PyMC3, and we sampled from the posterior distribution using NUTS with automatic differentiation variational inference initialization. We ran four Markov chain Monte Carlo 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.

From the posterior distribution, we estimated (a) the peak efficiency, namely the point at which observers' recognition performance reaches its maximum before

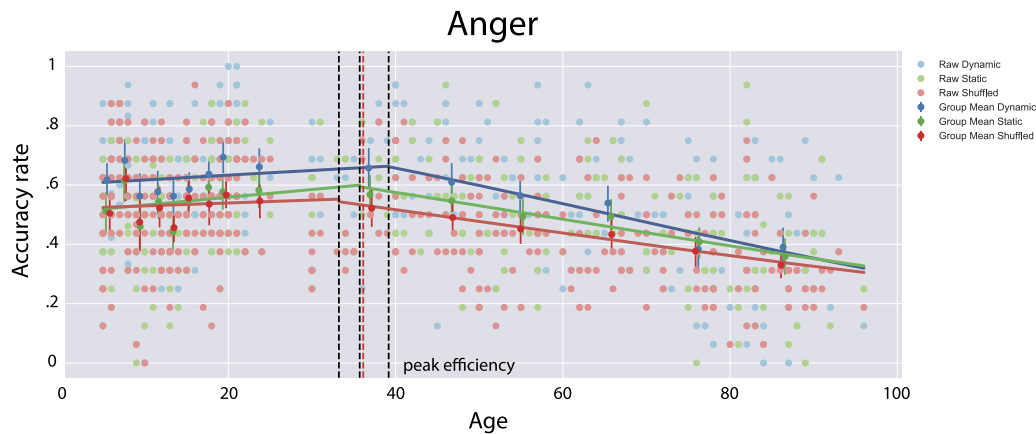


Figure 5. Anger. The posterior model fit (solid line) for the expression of anger with the individual performance (scatter plot) and the group average performance (dots with error bars) is given here. The overall peak efficiency is shown as the red vertical dashed line, and the condition-specific peak efficiencies are represented by the black dashed lines.

declining; (b) the steepness of increase and decrease in recognition abilities; (c) differences in the steepness of increase and decrease between different conditions (e.g., dynamic vs. static); and (d) the overall processing advantage of the dynamic over the static and the shuffled stimuli. By performing statistical inference directly on the full posterior distribution, we were able to properly quantify the dynamic stimuli effects and their associated uncertainty. In fact, the slopes of the linear relationship before and after the peak efficiency could provide estimations of the developmental trajectory and can be useful to make predictions on the performance of any new participant from any age. Here, however, we were only interested in comparing the slope difference between dynamic and static faces. A conceptual representation of the model is provided in Figure 3.

Results

The group average categorization performance for each condition is presented in Figure 4. The nonlinear relationship between age and recognition ability is clearly demonstrated with differences among conditions clearly visible for some expressions. When the model returns a concave pattern, we refer to the break point as a peak efficiency. This value relates to the point at which recognition performance reaches its apex, also relating to the age at which observers are the most efficient.

For the Bayesian modeling, trace plots, posterior distributions for the key parameters in the model, contrasts of interest, and full numerical reports of the parameter estimations are available in the supplementary results. Below, we report the key findings of the step-linear model.

Anger

The posterior model fit for the raw data is shown in Figure 5. By sampling the full posterior distribution, we estimated that the overall recognition ability for the expression of anger peaks at age 36.13 [22.23, 51.21], (bracket shows 95% highest posterior density interval). The posterior expectation of the age at which observers are the most efficient is given as follows: dynamic 39.17 [31.03, 46.84], static 35.70 [23.59, 46.41], and shuffled 33.22 [18.13, 49.62]. The overall recognition ability of anger at peak efficiency is 0.605 [0.350, 0.872], and the average peak accuracy for each condition is given as follows: dynamic 0.660 [0.627, 0.696], static 0.592 [0.554, 0.628], and shuffled 0.539 [0.489, 0.587]. On average, participants showed better performance in the dynamic condition as compared to the static and the shuffled conditions, both before (dynamic–static: 0.075 [0.042, 0.106], dynamic–shuffled: 0.096 [0.060, 0.132]) and after (dynamic–static: 0.043 [0.012, 0.072], dynamic–shuffled: 0.085 [0.051, 0.118]) peak efficiency. In contrast, the difference between the static and shuffled conditions is quite small (shuffled–static before peak efficiency: -0.021 [-0.061 , 0.019], after peak efficiency: -0.042 [-0.070 , -0.013]). The slopes of the step-linear functions are the following: dynamic 0.0069 [-0.0011 , 0.0156], static 0.0107 [0.0003, 0.0239], shuffled 0.0040 [-0.0072 , 0.0180] before peak efficiency and dynamic -0.0251 [-0.0300 , -0.0201], static -0.0183 [-0.0236 , -0.0123], shuffled -0.0159 [-0.0211 , -0.0108] after peak efficiency. Moreover, the differences of the slope across different conditions are mostly negligible; most of the posterior contrasts are distributed around zero with the exception of the contrast: dynamic–shuffled: -0.0092 [-0.0162 , -0.0020] after peak efficiency (Figure 5).

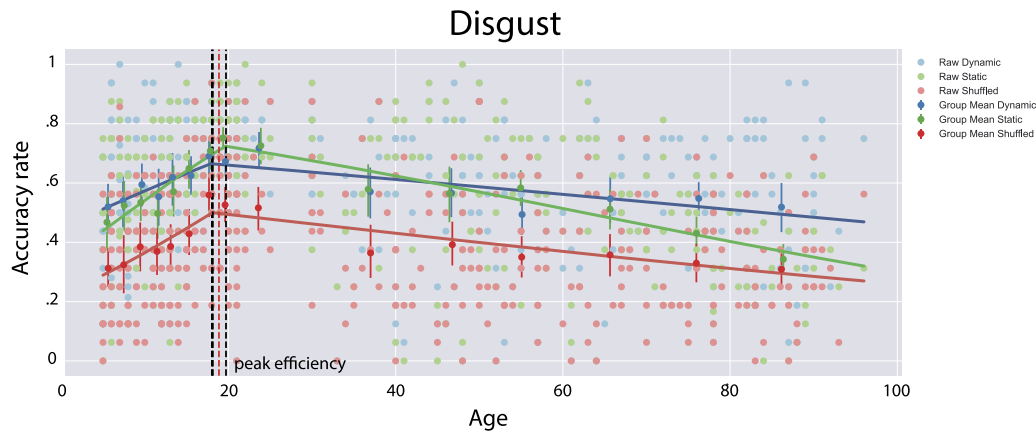


Figure 6. Disgust. The posterior model fit (solid line) of the expression of disgust with the individual performance (scatter plot) and the group average performance (dots with error bars) is given here. The overall peak efficiency is shown as the red vertical dashed line, and the condition-specific peak efficiencies are represented by the black dashed lines.

Disgust

The overall recognition ability for the expression of disgust peaks at age 18.87. The posterior expectation of the age at which observers are the most efficient is given as follows: dynamic 18.02 [15.38, 20.75], static 19.71 [18.19, 21.38], and shuffled 18.14 [16.55, 19.67]. The overall recognition ability of the expression of disgust at peak efficiency is 0.644 [0.380, 0.920], and the average peak accuracy for each condition is the following: dynamic 0.665 [0.644, 0.687], static 0.724 [0.702, 0.744], and shuffled 0.500 [0.475, 0.522]. On average, participants showed better performance in the dynamic and static conditions as compared to the shuffled condition, both before (dynamic–shuffled: 0.199 [0.160, 0.240], static–shuffled: 0.183 [0.146, 0.217]) and after (dynamic–shuffled: 0.191 [0.170, 0.214], static–shuffled: 0.152 [0.125, 0.177]) peak efficiency. The difference between the dynamic and the static conditions is quite small before peak efficiency (dynamic–static: -0.016 [-0.037 , 0.072]); it is, however, substantial after peak efficiency (0.040 [0.013 , 0.069]). The slopes of the step-linear functions are the following: dynamic 0.0494 [0.0318, 0.0680], static 0.0824 [0.0671, 0.0993], shuffled 0.0686 [0.0511, 0.0860] before peak efficiency and dynamic -0.0104 [-0.0130 , -0.0078], static -0.0225 [-0.0254 , -0.0196], shuffled -0.0128 [-0.0152 , -0.0102] after peak efficiency. Moreover, the slopes of the static condition are steeper than the ones in the dynamic and shuffled conditions. The contrasts of the slopes before peak efficiency are given as follows: dynamic–static: -0.0330 [-0.0577 , -0.0096], shuffled–static: -0.0138 [-0.0360 , 0.0123]; and the contrasts of the slopes after peak efficiency are the following: dynamic–static: 0.0121 [0.0084, 0.0161], shuffled–static: 0.0097 [0.0059, 0.0135] (Figure 6).

Fear

The overall recognition ability of the expression of fear peaks at around age 20.83. The posterior expectation of the age at which observers are the most efficient is given as follows: dynamic 20.87 [18.71, 23.18], static 19.72 [17.63, 21.43], and shuffled 21.79 [20.17, 23.51]. The overall recognition ability of fear at peak efficiency is 0.446 [0.168, 0.697]; the average peak accuracy for each condition is the following: dynamic 0.399 [0.372, 0.430], static 0.416 [0.390, 0.443], and shuffled 0.526 [0.494, 0.557]. On average, participants showed better performance in the shuffled condition compared to the other two conditions, both before (shuffled–dynamic: 0.056 [0.013, 0.099], shuffled–static: 0.055 [0.024, 0.083]) and after (shuffled–dynamic: 0.094 [0.071, 0.118], shuffled–static: 0.113 [0.089, 0.135]) peak efficiency; however, the difference between the dynamic and static conditions is quite small (dynamic–static before peak efficiency: -0.002 [-0.043 , 0.033], after peak efficiency: 0.018 [-0.003 , 0.040]). The slopes of all conditions are comparable: dynamic 0.0706 [0.0546, 0.0867], static 0.0927 [0.0747, 0.1139], shuffled 0.0913 [0.0764, 0.1056] before peak efficiency and dynamic -0.0151 [-0.0186 , -0.0115], static -0.0189 [-0.0224 , -0.0157], shuffled -0.0176 [-0.0213 , -0.0141] after peak efficiency. The maximum contrasts of the slopes before peak efficiency is given as follows: dynamic–static: -0.022 [-0.0473 , 0.0022], and the maximum contrasts of the slopes after peak efficiency is the following: dynamic–static: 0.0038 [-0.0010 , 0.0086] (Figure 7).

Happiness

Unlike for the other facial expressions, the overall recognition ability for the expression of happiness is

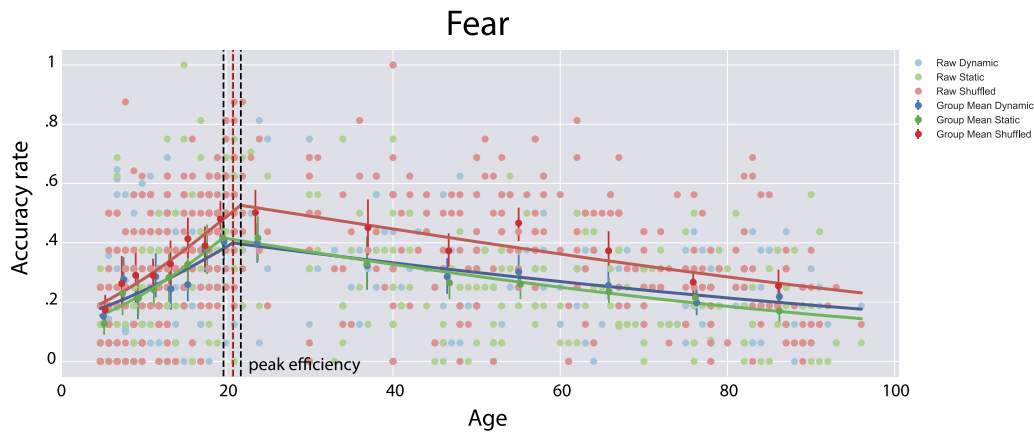


Figure 7. Fear. The posterior model fit (solid line) of the expression of fear with the individual performance (scatter plot) and the group average performance (dots with error bars) is presented here. The overall peak efficiency is shown as the red vertical dashed line, and the condition-specific peak efficiencies are represented by the black dashed lines.

near ceiling at a very young age, declining slowly throughout the life span. Therefore, the age of the peak efficiency could not be identified for this facial expression. Nonetheless, our model identifies a break point at around age 57.98 with a large uncertainty. Importantly, the accuracy rate estimated at this break point is not the apex in recognition performance, but rather the start of the decline (i.e., the model did not return a concave pattern). The posterior expectation of the age at this break point is given as follows: dynamic 61.27 [24.18, 93.42], static 50.25 [35.90, 62.01], and shuffled 62.22 [21.95, 81.00]. The overall recognition ability of this expression at the break point is 0.855 [0.664, 0.999] with dynamic 0.895 [0.819, 0.972], static 0.898 [0.868, 0.931], and shuffled 0.660 [0.541, 0.896]. Overall, participants performed better in the dynamic condition as compared to the static and shuffled conditions, both before (dynamic–static: 0.040 [0.027, 0.054], dynamic–shuffled: 0.106 [0.063, 0.143]) and after

(dynamic–static: 0.060 [0.029, 0.095], dynamic–shuffled: 0.269 [0.203, 0.319]) the break point. Participants also performed better in the static than in the shuffled condition (static–shuffled before the break point: 0.066 [0.018, 0.105] and after the break point: 0.209 [0.156, 0.255]). The slopes of the step-linear functions are given as follows: dynamic -0.0248 [-0.0335 , -0.0121], static -0.0078 [-0.0172 , 0.0007], shuffled -0.0306 [-0.0375 , -0.0163] before the break point and dynamic -0.0239 [-0.0387 , -0.0023], static -0.0298 [-0.0381 , -0.0217], shuffled -0.0175 [-0.0355 , 0.0026] after the break point. The differences of the slopes across the different conditions are mostly negligible. Most of the posterior contrasts are distributed around zero with the largest contrasts being the following before the break point: dynamic–static: -0.0170 [-0.0297 , -0.0036] and shuffled–static: -0.0228 [-0.0353 , 0.0087] (Figure 8).

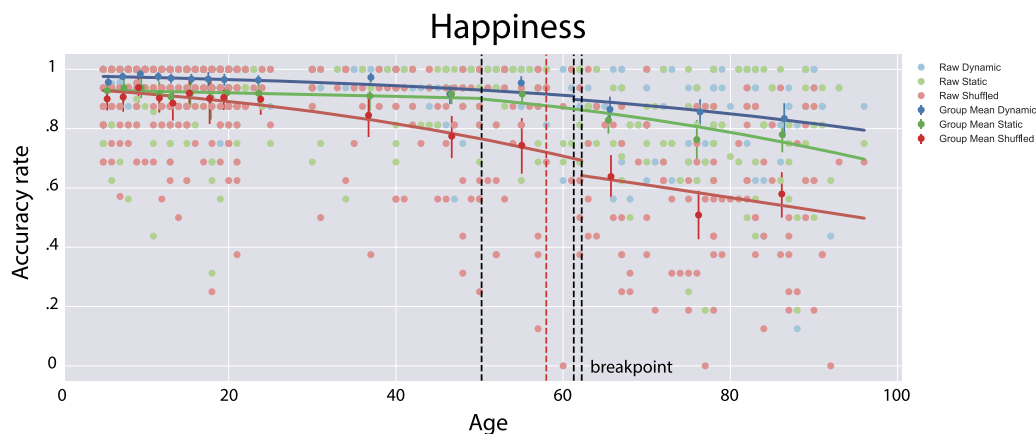


Figure 8. Happiness. The posterior model fit (solid line) of the expression of happiness with the individual performance (scatter plot) and the group average performance (dots with error bars) is presented here. The overall break point is shown as the red vertical dashed line, and the condition-specific break points are represented by the black dashed lines.



Figure 9. Sadness. The posterior model fit (solid line) of the expression of sadness with the individual performance (scatter plot) and the group average performance (dots with error bars). The overall peak efficiency is shown as the red vertical dashed line, and the condition-specific peak efficiencies are represented by the black dashed lines.

Sadness

The overall recognition ability for the expression of sadness peaks at the age of 28.96. The posterior expectation of the age at which observers are the most efficient is given as follows: dynamic 27.87 [21.44, 33.90], static 28.12 [23.50, 32.62], and shuffled 30.52 [26.05, 34.74]. The overall recognition ability of sadness at peak efficiency is 0.638 [0.408, 0.888]; the average peak accuracy for each condition is the following: dynamic 0.605 [0.572, 0.636], static 0.631 [0.602, 0.660], and shuffled 0.653 [0.622, 0.681]. The categorization accuracy rates of all conditions are comparable both before and after peak efficiency. The maximum contrasts of the average performance before peak efficiency is given as follows: shuffled–static 0.0277 [−0.0011, 0.0561]; the maximum contrasts of the average performance after peak efficiency is the following: shuffled–static 0.0097 [−0.0156, 0.0362]. Similarly, all conditions show comparable slopes: dynamic 0.0075 [−0.0031, 0.0181], static 0.0235 [0.0117, 0.0347], shuffled 0.0179 [0.0079, 0.0285] before peak efficiency and dynamic −0.0223 [−0.0262, −0.0186], static −0.0265 [−0.0306, −0.0223], shuffled −0.0302 [−0.0346, −0.0261] after peak efficiency. The maximum contrast between slopes before peak efficiency is the following: dynamic–static: −0.0160 [−0.0320, −0.0010]; the maximum contrast after peak efficiency is given as follows: dynamic–shuffled: 0.0079 [0.0022, 0.0137] (Figure 9).

Surprise

The overall recognition ability of surprise peaks at age 22.47. The posterior expectation of the age at which observers are the most efficient is given as follows: dynamic 23.55 [20.52, 26.91], static 24.30 [20.36, 28.26],

and shuffled 19.34 [17.34, 21.76]. The overall recognition ability of surprise at peak efficiency is 0.692 [0.466, 0.953]; the average peak accuracy for each condition is dynamic 0.758 [0.735, 0.783], static 0.700 [0.673, 0.725], and shuffled 0.575 [0.552, 0.599]. On average, participants showed the best performance in the dynamic condition, and the worst in the shuffled condition. The results were the following: dynamic–static: 0.075 [0.048, 0.101], static–shuffled: 0.093 [0.058, 0.123] before peak efficiency and dynamic–static: 0.107 [0.082, 0.133], static–shuffled: 0.172 [0.146, 0.195] after peak efficiency. The slopes of the step-linear functions are dynamic 0.0442 [0.0315, 0.0565], static 0.0442 [0.0311, 0.0577], shuffled 0.0530 [0.0381, 0.0692] before peak efficiency and dynamic −0.0126 [−0.0164, −0.0092], static −0.0175 [−0.0213, −0.0133], shuffled −0.0190 [−0.0220, 0.0163] after peak efficiency. The slope between age and accuracy is similar across all conditions before peak efficiency, whereas after peak efficiency, the dynamic condition shows the most gradual slope: static–dynamic: −0.0048 [−0.0104, 0.0004], shuffled–dynamic: −0.0064 [−0.0109, −0.0018] (Figure 10).

To highlight the key results of our study, we illustrated the results of the previous statistical analyses comparing the dynamic and static tasks in Figure 11. The posterior estimation of peak efficiency ages in the recognition of all the expressions except happiness are summarized in Figure 11A. The estimated age is the youngest for disgust and fear, whereas the oldest for anger (which is also the expression driving the largest uncertainty in the estimation). Moreover, the dynamic and static conditions do not significantly modulate this effect.

The differences in recognition between the dynamic and static conditions are summarized in Figure 11B, which shows performance both before and after the estimated peak efficiency. As reported below, the recognition performance of the facial expression of

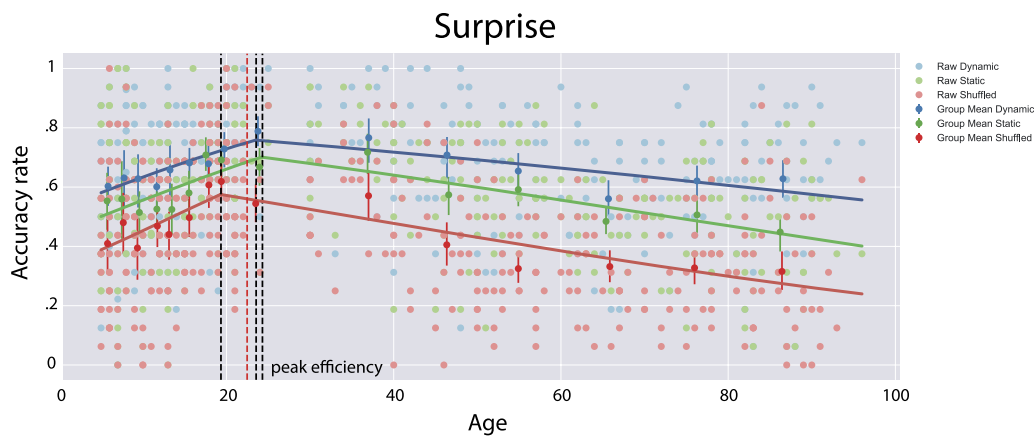


Figure 10. Surprise. The posterior model fit (solid line) for the expression of surprise with the individual performance (scatter plot) and the group average performance (dots with error bars) are given here. The overall peak efficiency is shown as the red vertical dashed line, and the condition-specific peak efficiencies are represented by the black dashed lines.

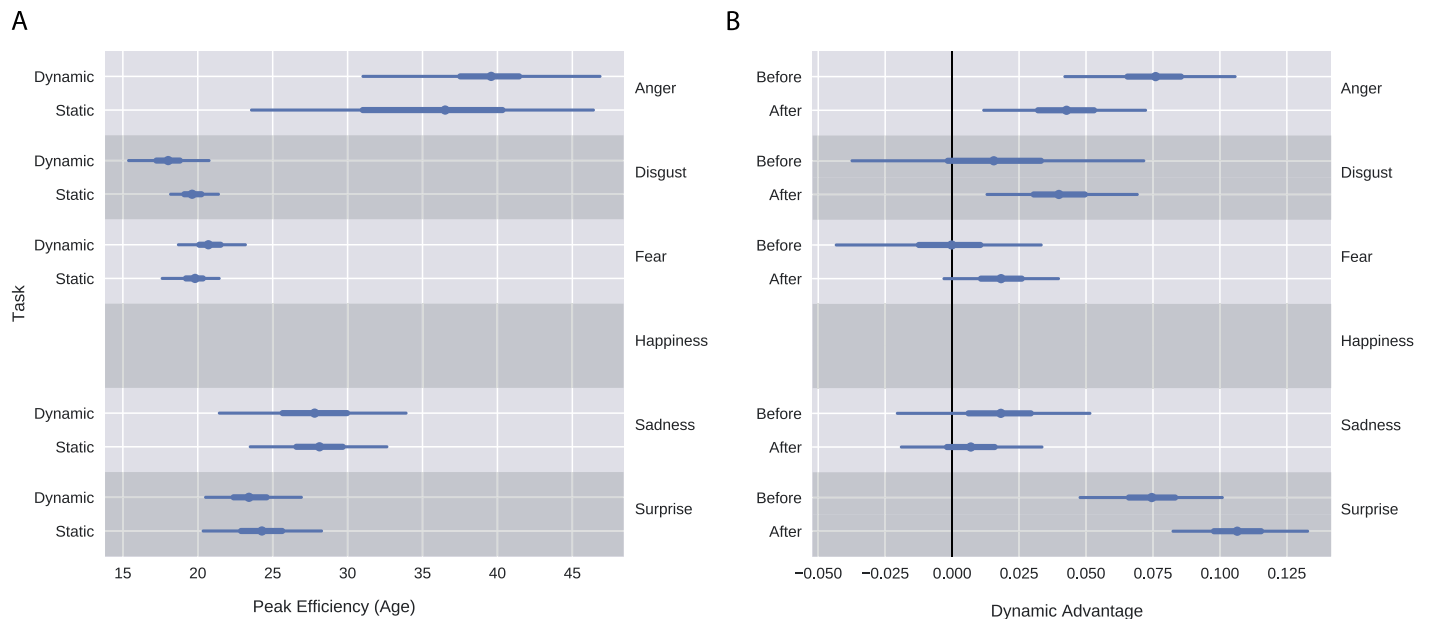


Figure 11. Summary of the key findings. (A) The posterior estimation of peak efficiency ages in the recognition of different facial expressions of emotion. (B) The posterior estimation of the dynamic advantage *before* and *after* the peak efficiency age (represented by the average difference between the correct categorization of dynamic and static facial expressions). The recognition of the happy expression could not be identified as the performance for this facial expression was already at ceiling in the early age we tested. The dots show the posterior expectation, the bold horizontal line shows the 50% highest posterior density, and the thin horizontal line shows the 95% highest posterior density. Nonoverlapping lines indicate a significant difference between two conditions.

surprise shows the largest dynamic advantage, whereas fear is, on average, similar in both conditions.

Discussion

Our results present a fine-grained developmental tracking of human observers' ability to recognize the six basic emotions when presented with varying temporal

properties: dynamic, static, and shuffled. Previous studies in the literature examined expression recognition by using *arbitrary* age groups: 10-year bins (Williams et al., 2009), stages of life (Horning, Cornwell, & Davis, 2012), or largely different age groups (e.g., 18–30, 58–70, Calder et al., 2003) while revealing either expression-recognition improvement (Rodger et al., 2015) or decline (Calder et al., 2003; MacPherson et al., 2002; Malatesta, Izard, Culver, & Nicolich, 1987; Moreno, Borod, Welkowitz, & Alpert, 1993; Ruffman et al., 2008;

Sullivan & Ruffman, 2004a). In contrast, our approach innovates by estimating the continuous developmental trajectory of facial expression recognition (from increase to decline) by considering age as a *continuum*, ranging from 5 to 96 years.

Using a Bayesian approach, we estimated for each condition and each expression individually the associated uncertainty and (a) the peak efficiency, namely the point at which observers' recognition performance reaches its maximum before declining; (b) the steepness of increase and decrease in recognition abilities; (c) differences in the steepness of increase and decrease between different conditions (e.g., dynamic vs. static); and (d) the overall processing advantage of the dynamic over the static and the shuffled stimuli. We now discuss, in turn, each of these findings and their implications.

Recognition trajectory across development: Increase, peak efficiency, and decrease

Our findings reveal unique developmental profiles and peak efficiency for the static, dynamic, and shuffled versions of each individual expression. Herein, we focus on the dynamic and static trajectories and the differences between both conditions in a more detailed manner (i.e., static and dynamic). The results of the shuffled condition are briefly considered at the end of the discussion.

Efficiency: Increase

In both static and dynamic conditions, the sharpest rises in accuracy were observed for fear followed by disgust and, to a lesser extent, surprise. These findings mirror the results of a previous developmental study that investigated the effects of age on the development of emotion processing in children, revealing that increasing age produced significant improvements in the recognition of fear and disgust (Herba & Phillips, 2004). We observed a more gradual increase for sadness and anger but only in the static condition. Finally, we did not observe any increase for the expression of happiness regardless of the experimental condition.

The steepest increase observed for fear might be accounted for by the very low recognition rates observed for this expression in young children, reaching only 13% in the 5–6 age group in the static condition (15% in the dynamic condition; Supplementary Figure S1A). The expression of fear has been regularly reported in developmental (Herba & Phillips, 2004; Rodger et al., 2015; Widen, 2013), neuropsychological (Adolphs et al., 2003; Richoz et al., 2015), and behavioral studies (Calder et al., 2003) as being the

most difficult expression to effectively recognize among all the expressions—a difficulty that is puzzling considering the evolutionary importance of an adequate and rapid categorization of this expression for survival. Importantly, however, a poor performance for the recognition of some expressions does not necessarily mean that these expressions are not detected as recent studies have shown a dissociation between these two processes (Smith & Rossit, 2018; Sweeny, Suzuki, Grabowecy, & Paller, 2013). For example, Smith and Rossit (2018) have shown that the emotional expression of fear is better detected than recognized. Also, among the basic expressions, fear is probably the one that transmits the strongest multisensory perceptual cues. Multisensory and contextual information, such as environmental threats, may, therefore, play a crucial role in the decoding of this expression and be essential for an adequate categorization.

Consistent with our findings, fear has been observed to display a sharp increase in some prior developmental studies (Herba, Landau, Russell, Ecker, & Phillips, 2006; Vicari, Reilly, Pasqualetti, Vizzotto, & Caltagirone, 2000), and other studies have revealed more gradual improvements (Gao & Maurer, 2009; Thomas, De Bellis, Graham, & LaBar, 2007) or stable, albeit low task performance from early childhood to adulthood (Rodger et al., 2015). Differences across studies may be attributed to methodological considerations and task differences as recognition rates have been proven to be task dependent (e.g., Montiroso, Peverelli, Frigerio, Crespi, & Borgatti, 2010; Vicari et al., 2000) with performance variations occurring even within the same study when the task is changed (Vicari et al., 2000). Importantly, the findings reported here provide further evidence that the recognition of fear has a special status within the framework of facial-expression recognition (Richoz et al., 2015; Rodger et al., 2015).

Disgust also showed a steep increase in recognition accuracy, following a similar trajectory as fear. In line with our findings, steep improvements from childhood to adulthood were previously observed for disgust in a study by Rodger et al. (2015), which measured the quantity of information necessary for an observer to accurately recognize facial expressions, as well as in earlier studies that investigated expression recognition with matching (Herba et al., 2006) or labeling tasks (Vicari et al., 2000). As mentioned by Vicari et al. (2000), the steep improvement observed for disgust in children aged 5 to 10 may occur owing to the greater lexico-semantic abilities in older children. It might also be plausible that the very distinctive facial configurations of disgust convey signals about potentially contaminated food. These signals are crucial from an evolutionary perspective and, hence, the need to rapidly improve in the detection of this expression in order to stay away from harmful substances.

Finally, our findings also revealed a sharp increase for surprise in both the static and dynamic conditions. Interestingly, the expression of surprise was already well recognized in very young children aged 5 to 6 with recognition rates of 60% for the dynamic stimuli (55% for the static images; Supplementary Figure S1A). High recognition rates in young children would rather accord with a more gradual developmental trajectory as suggested by prior research that investigated the recognition of surprise from 5 up to 18 years of age (Rodger et al., 2015). Interestingly, however, the sharp increase observed for surprise in the current study may be accounted for by the very high recognition rates observed for this expression in participants above the age of 18, reaching up to 79% in the 21–30 age group for the dynamic stimuli (67% for the static images; Supplementary Figure S1C).

We also observed a gradual increase for anger and sadness although only in the static condition. These findings are generally consistent with previous reports (Herba et al., 2006; Rodger et al., 2015, for sad; Vicari et al., 2000). In the dynamic condition, children aged 5 to 6 were nearly as effective in recognizing anger (62%; Supplementary Figure S1A) as young adults aged 17–18 (64%; Supplementary Figure S1B), 21–30 (66%; Supplementary Figure S1C), or 31–40 (66%; Supplementary Figure S1D). The same pattern was observed for sadness with identical recognition rates for young children aged 5–6 (56%) and young adults in the 17–18 age group.

Finally, our results did not reveal an increase for happiness in either condition with task performance remaining stable over ages and with our peak efficiency revealing the peak of the decline. The absence of improvement observed for happiness may be explained by the very high recognition accuracy already found in young children for this expression, which leaves little scope for improvement. Our findings for happiness are consistent with previous studies that revealed that children as young as 5 years of age recognize the expression of happiness just as effectively as adults (Gao & Maurer, 2009; Gross & Ballif, 1991; Herba & Phillips, 2004) even when the presentation time was as fast as 500 ms (Rodger et al., 2015). In order to capture the increase in recognition performance for happiness, we should have started earlier, under 5 years of age, or adopt alternative approaches. Limiting the available information by degrading the signal (Rodger et al., 2015), controlling for spatial frequency use (Gao & Maurer, 2011), or modifying the intensity of facial expressions (Gao & Maurer, 2009, 2010; Rodger, Lao, & Caldara, 2018) all represent alternative techniques to potentially reveal the improvement in the recognition of the emotional expression of happiness. Therefore, future studies using those approaches represent a promising route to assess and eventually reveal

developmental differences in the trajectory of the decoding of happiness.

Additionally, it is worth noting that our findings revealed differences in the steepness of increase between the static and dynamic conditions for the expressions of disgust and sadness, the increase being steeper with the static stimuli. These findings might be accounted for by the low recognition rates found for the static images of disgust and sadness in very young children. An exposure to static images of disgust and sadness is rather uncommon in everyday life, particularly for young children, whereas an exposure to the dynamic versions of these expressions might be more frequent for children when their schoolfellows or siblings dislike some particular food they have to eat (disgust) or when they cry or express their sorrow (sadness).

Peak efficiency

The data-driven identification of the peak efficiency, the point at which observers' recognition performance reaches its optimum before declining, revealed a series of novel interesting findings. To the best of our knowledge, this is the first study that has effectively isolated the age at which observers are the most efficient for the recognition of the basic facial expressions of emotion across the life span. We observed the earliest peak efficiencies for both the static and dynamic expressions of disgust (18–20 years) and fear (19–21 years) in young adults, followed by surprise (23–25 years) and sadness (27–28 years). Peak efficiency for the static expression of anger was found at 35 years of age, whereas the recognition of its dynamic version was reached at 39 years. The latest break point that emerged from our data was observed for the dynamic expression of happiness at around 61 years of age (50 years for the static version). It is worth noting that the break points found for each expression were nearly the same in both conditions with the exception of anger and happiness, which reached their break points later with dynamic expressions.

There are two explanations for the very early peak efficiencies found for fear and disgust. First, as mentioned above, from an evolutionary perspective, these two expressions convey important signals about potential dangers or harmful substances, both important for survival. Disgust and fear can, therefore, be expected to reach their peak rapidly in order to ensure survival. Second, for fear and disgust, the point in time at which the peak efficiency emerges may be driven by the inherent properties of those expressions. Stimuli that are difficult to recognize for young people might be even more difficult for elderly people as difficult tasks are likely to be more sensitive to cognitive decline (Calder et al., 2003; Ruffman et al., 2008). Changes in

the slope of the lines may, therefore, be expected to occur earlier with difficult tasks. We examined response biases for each expression, computing confusion matrices across different age groups (see Supplementary Figure S1A through E). The confusion matrices found for fear and disgust indeed revealed that these two expressions were particularly difficult for our observers to identify. Disgust was commonly confused with anger with confusion rates ranging up to 28% in the 5–6 age group for the dynamic stimuli (Supplementary Figure S1A) and 20% for the 71–80 age group (Supplementary Figure S1E). Previous studies also reported marked confusion between disgusted and angry faces, which were interpreted as a general bias toward angry faces (Recio et al., 2013). Such a bias could explain the stable and high recognition rates found in the current research for angry faces from childhood onward. Other confusion was observed between fear and surprise. In line with previous studies (Rodger et al., 2015), fear was found to be the most frequently confounded expression among all age groups with confusion rates reaching up to 53% for the dynamic expression of surprise in the 5–6 age group (Supplementary Figure S1A) or even 63% in the 71–80 age group (Supplementary Figure S1E). As mentioned by Calder et al. (2003), age-related cognitive decline may reinforce these confusions due to perceptual or conceptual difficulties (i.e., fear and surprise are conceptually very close and share facial signals that are morphologically similar; Delis et al., 2016). Note also that the reverse confusion was much less common. When presented with surprise, the confusion rates observed for fear reached only 3% for the dynamic stimuli (4.6% for the static expressions) in the 5–6 age group and 5% (3.3% for the static expressions) in elderly people aged 71–80.

Interestingly, our findings also revealed a later emergence of the peak efficiency for anger compared to the other expressions. As mentioned before, recognition abilities for anger showed no increase in the dynamic condition, task performance being already high in young children, and displayed only a slight increase for the static condition, recognition rates being also high in young children. A potential, although speculative, explanation for this observation may lie in the fact that we are daily exposed to the expression of anger, arguing with our partners, children, colleagues—an exposure that might postpone the recognition decrease of this expression and, therefore, the changes in the slope of the line.

Finally, the latest break point found for happiness may be accounted for by the ceiling effect found for this expression from childhood onward.

Altogether, this second set of findings offers novel insights into the development of human facial-expression recognition. As observed, facial-expression recog-

niton develops following *emotion-dependent* trajectories that do not necessarily all reach their peak efficiency in early adulthood as predicted by previous studies (Calder et al., 2003; De Sonneville et al., 2002; Horning et al., 2012; Williams et al., 2009). The optimal level of task performance can indeed be reached at a very late point in development, depending also on the very nature of the diagnostic information of the facial expression and its temporal properties and evolutionary value.

Efficiency: Decrease

Finally, we observed differences in the steepness of decrease in recognition performance across emotions and conditions. In the *dynamic* condition, the steepest decreases were observed for anger, happiness, and sadness and less severe decreases for fear and surprise. Disgust showed the least severe decrease in this condition. Different patterns were observed in the *static* condition, the steepest decline being for happiness followed by sadness and disgust. Less severe decreases were found for fear and anger, whereas the least severe decrease was observed for the expression of surprise. Similarly to the differences in the steepness of increase observed between static and dynamic conditions, differences in the steepness of decrease were observed between both conditions for the expression of disgust. The recognition of the static expression of disgust decreased from 51% to 34% between the ages of 61 and 70 and 81 and 90, whereas recognition accuracy of its dynamic version remained relatively stable (decrease from 55% to 52%; Supplementary Figure S1E).

This pattern of results posits that the recognition of facial expression declines over time, which is consistent with previous models of aging. These models suggest that age-related structural changes in different brain regions, particularly in frontal and temporal volumes as well as changes in neurotransmitters (Calder et al., 2003; Ruffman et al., 2008) might be responsible for older adults' impairment in the recognition of facial expression. For example, the amygdala, which plays a crucial role in the processing of fear and sadness (e.g., Adolphs et al., 2005; Yang et al., 2002), undergoes severe atrophy with age and becomes progressively less responsive to negative stimuli (De Winter et al., 2016; Mather et al., 2004; Ruffman et al., 2008). In contrast, the insula and basal ganglia, which underlie the processing of disgust, seem to be less vulnerable to aging as evidenced by the preserved ability to recognize this expression in older adults (Calder et al., 2003; Horning et al., 2012; Ruffman et al., 2008). Interestingly, our findings revealed the least severe decrease in recognition

performance for the dynamic expression of disgust. However, in contrast to previous studies that showed no reduction in the recognition of some expressions or even some improvements with increasing age (e.g., Calder et al., 2003), our findings revealed steep-to-moderate decreases for all the expressions even for disgust being usually preserved in elderly people (Calder et al., 2003; Horning et al., 2012; Ruffman et al., 2008).

Methodological considerations may be responsible for the differences observed between the current study and previous ones. Indeed, previous studies investigated facial expression recognition across groups of ages (Calder et al., 2003), stages of life (Horning et al., 2012), or decades (Williams et al., 2009), whereas our study investigated elderly people's ability to categorize emotions by considering age as a continuum. This methodological approach overcomes the problem of defining arbitrary age boundaries, which are routinely used in the literature to relate to critical developmental ages.

Furthermore, the variability in findings between the current research and previous neuropsychological and behavioral studies can be accounted for by the age ranges tested across the studies. For example, Calder et al. (2003) observed improved recognition abilities for disgust in their older adult age group, spanning from age 58 to 70 (mean age 65). In contrast, in our study, we tested participants up to the age of 96, giving rise to the possibility that the decline for disgust appears at a later point in development. This assumption is in line with a previous study that showed a decrease in the recognition of disgust in elderly people aged 80 to 91 (Williams et al., 2009).

Additionally, the stimuli used across the different studies might also have impacted expression-recognition performance. In the current study, we used a specific database of emotional expressions that are less prototypical than the Ekman and Friesen (1976) standard set of facial photographs used in previous research (Calder et al., 2003; McDowell, Harrison, & Demaree, 1994; Sullivan & Ruffman, 2004b). Moreover, in contrast to previous reports that used only static images displaying the apex or the highest state of an emotional expression (Calder et al., 2003; Ruffman et al., 2008), we tested facial expression recognition with static, dynamic, and shuffled stimuli. Importantly, our stimuli were controlled for the amount of low-level discriminative information carried over time. In other words, the quantity of low-level information carried by our static, dynamic, and shuffled stimuli was identical across conditions and tasks (Gold et al., 2013). In line with previous studies (Krendl & Ambady, 2010; Sze et al., 2012), we found that elderly people were impaired in recognizing static but not dynamic expressions. However, in contrast to

the findings reported by Krendl and Ambady (2010), we also observed steep-to-moderate declines for all the expressions even in the dynamic condition. However, in their study, participants were provided with additional aiding cues, such as body-related information or contextual cues, which might have facilitated expression recognition given that the perception of a particular expression is strongly influenced by the context in which it occurs (Barrett & Kensinger, 2010; Horning et al., 2012). For example, Aviezer et al. (2008) found more consistent recognition performance for fear when person-related or contextual information was provided to the participants.

Finally, the divergence between our findings and those of previous research may also be due to the small number of trials presented (Horning et al., 2012; Moreno et al., 1993) as well as the differences in the settings used with some studies relying on laboratory settings (Calder et al., 2003; Horning et al., 2012) and others on online tasks (Williams et al., 2009).

Static versus dynamic expressions

A dynamic advantage before peak efficiency

Our findings revealed a dynamic face advantage for the recognition of anger, surprise, and happiness before peak efficiency. These results are inconsistent with previous developmental studies, which revealed that dynamic presentations did not increase children's recognition performance (Nelson et al., 2013; Nelson & Russell, 2011; Widen & Russell, 2015), and with some experiments showing even an overall advantage for static expressions (Nelson & Russell, 2011, study 1; Widen & Russell, 2015). Such advantage also differs from the results reported by previous studies in young and healthy adults (e.g., Christie & Bruce, 1998; Jiang et al., 2014; Kätsyri & Sams, 2008), showing that the recognition of facial expressions is not facilitated by the dynamic information provided by moving faces.

The lack of consistency between these studies and the present work may be accounted for by methodological factors. For instance, in some of the aforementioned developmental studies, only a single actor was selected to record the facial expressions (Nelson et al., 2013; Nelson & Russell, 2011), raising the possibility that the results found could be biased by the acting performance. Compared to the Ekman and Friesen (1976) standard set of facial expressions, the expressions of the single actor used in the study by Nelson and Russell (2011) were indeed more readily labeled as correct by adults as they were perceived as clearer and more intense. Asking children to catego-

size facial expressions of a single actor in their dynamic and static forms might also have impacted their recognition performance because they might have been more likely to choose the same label in both the conditions by using a picture-matching strategy. In addition, compared to the current research in which children were asked to choose the correct answer among six possibilities, previous developmental studies used free labeling as a measure of recognition (Nelson & Russell, 2011; Widen & Russell, 2015), which raises the possibility that vocabulary performance rather than children's true ability to understand the emotions of others were tested.

Importantly, most developmental studies that revealed an overall static advantage for facial expression recognition (Nelson & Russell, 2011; Widen & Russell, 2015) directly compared children's performance for static expressions to their scores with dynamic expressions. In most cases, these direct comparisons can be problematic because they make it difficult to determine whether increased recognition rates are caused by psychological or physical factors (Gold et al., 2013). For instance, Nelson and Russell (2011) and Widen and Russell (2015) created their static images by presenting a single frame of the highest amplitude of the dynamic sequences, a procedure that might have created "optimal" static images. The overall static advantage found in their research may be due to an increased quantity of discriminative information provided by the stimuli rather than an enhanced psychological ability to perceive the static expressions. In order to control for this general confounding of physical and psychological factors, we decided to rely on a database of stimuli created by Gold et al. (2013), who controlled for the amount of low-level information carried by their stimuli over time by carefully dissociating these two factors with the use of a psychophysical approach. Compared to previous studies (Nelson et al., 2013; Nelson & Russell, 2011; Widen & Russell, 2015), our results offer a more reliable view and a better understanding of the way in which temporal properties influence facial expression recognition from childhood onward.

A dynamic advantage after peak efficiency

Our results revealed processing benefits of dynamic stimuli after peak efficiency for all the expressions with the exception of sadness and fear. Interestingly, our data evidenced that these results were driven by a suboptimal performance for the recognition of static expressions in elderly people rather than increased abilities to recognize dynamic expressions (see Supplementary Table S1 for the example of surprise).

In everyday life, facial expressions are dynamic events that unfold over time in some particular ways, representing a richer and more valid approach to study facial expression recognition. Previous fMRI studies have also suggested that different neural substrates underlie the processing of dynamic and static expressions (e.g., Johnston et al., 2013; Kessler et al., 2011; LaBar, Crupain, Voyvodic, & McCarthy, 2003; Paulmann et al., 2009; Sato, Kochiyama, Yoshikawa, Naito, & Matsumura, 2004; Schultz & Pilz, 2009; Trautmann et al., 2009). Dynamic faces have been found to selectively elicit higher neural responses in the posterior superior temporal sulcus, in the anterior superior temporal sulcus, and in the inferior frontal gyrus (Bernstein, Erez, Blank, & Yovel, 2017; Fox, Iaria, & Barton, 2009; Pitcher, Dilks, Saxe, Triantafyllou, & Kanwisher, 2011). A very recent fMRI study that used multivoxel pattern analysis revealed that dynamic expressions were associated with increased activation in both face-selective and motion-selective areas as well as higher categorization accuracies compared to static expressions (Liang et al., 2017). Given that dynamic faces elicit higher neural responses (Bernstein et al., 2017; Fox et al., 2009; Pitcher et al., 2011) and cause the activation of a wider network of regions in the brain (Arsalidou, Morris, & Taylor, 2011; Liang et al., 2017), their decoding may be less vulnerable to age-related degeneration compared to the decoding of static images.

In contrast, suboptimal performance for static stimuli could be explained by age-related structural changes in brain regions responsible for the processing of static emotional expressions. For example, De Winter et al. (2016) recently evidenced that age-induced atrophy to the amygdala of patients with frontotemporal dementia affected emotion processing in distant face-selective areas. More specifically, their findings evidenced a positive correlation between gray matter volume in the left amygdala and emotion-related brain activity in the fusiform face area, a core region in the face-processing network involved in the decoding of static stimuli (Pitcher et al., 2011) and emotional expressions (Ganel, Valyear, Goshen-Gottstein, & Goodale, 2005; Xu & Biederman, 2010).

Dynamic faces also include information that cannot be completely rendered by static images, forcing the observers to shift their attention to different facial features. Multiple shifts on different facial areas are likely to benefit expression recognition given that the facial signals critical for the recognition of emotional expressions can be found throughout the face. In clinical conditions or normal aging populations, when slower or suboptimal processing takes place, dynamic stimuli may provide additional cues, attracting and holding attention as well as enhancing motor simulations (Rymarczyk, Biele, Grabowska, & Majczynski,

2011; Sato, Fujimura, & Suzuki, 2008; Sato & Yoshikawa, 2007). The increased attention inherently elicited by moving faces may compensate for the apparent age-related deficits found in elderly populations on expression-recognition tasks with static images. Dynamic face stimuli may naturally drive the focus of attention toward the diagnostic information in a bottom-up fashion (i.e., the mouth for surprise), whereas static face images require the observers to move toward those features based on top-down internal representations.

It is also important to note that the advantage we observed for dynamic expressions cannot simply be attributed to an overall larger amount of discriminative information carried by the dynamic stimuli. As reported above, the stimuli used in the current experiment were created by Gold et al. (2013), who used an ideal observer approach to effectively measure the amount of information provided by the stimuli. Gold et al.'s findings reveal that their dynamic stimuli did not offer more discriminative information to the observers as compared to their static images. Thus, the dynamic advantage for anger, disgust, happiness, and surprise observed in our participants is unlikely to be the result of physical factors. Rather, this dynamic advantage most probably comes from an adequate ability to use the available perceptual and diagnostic information.

We did not find a dynamic advantage for fear and sadness before or after peak efficiency. From a sociobiological perspective, the expression of fear is critical for human survival (LoBue, 2010), and it has been shown to potentiate early visual processing of perceptual events (e.g., Phelps, Ling, & Carrasco, 2006) and enhance attention (e.g., Carlson & Mujica-Parodi, 2015; Pourtois, de Gelder, Bol, & Crommelinck, 2005). Interestingly, using a single-trial repetition suppression approach, we very recently revealed that this expression boosts the early coding of individual faces regardless of attentional constraints (Turano et al., 2017). Similar findings were reported in a very recent developmental study that examined detection thresholds for happy and fearful faces presented with noise. The superior ability for detecting fearful faces was observed already in infants aged 3.5 months (Bayet et al., 2017). Our brain may be particularly tuned to recognize this expression regardless of its temporal properties. This assumption may explain the absence of a dynamic advantage for the decoding of this expression. However, enhanced processing of fear would also predict increased categorization performance, a prediction that is inconsistent with our findings. Our results indeed revealed very low recognition rates throughout the life span. Interestingly, in a recent study, Sweeny et al. (2013) evidenced that the detection of a fearful face dissociates from its categorization. Fearful faces were

very rapidly and accurately detected even when presented for only 10 ms, whereas their categorization rates were near chance level. Our brain might, therefore, be specially tuned to detect this expression independently from its categorization (see also Smith & Rossit, 2018). Also, as mentioned above, among all expressions, fear is probably the most powerful for transmitting multisensory information. Broader contextual information may, therefore, be necessary to reliably categorize it. This assumption is in line with an emerging literature that suggests that isolated facial signals may not be sufficient for observers to adequately perceive the emotions of fear and disgust and that additional information regarding the context in which the expression occurs is critical (e.g., Barrett & Kensinger, 2010).

The absence of a dynamic advantage for the processing of sadness is consistent with previous findings, which revealed that the expression of sadness is better recognized through static pictures (Bould et al., 2008; Recio et al., 2013; Widen & Russell, 2015) or when evolving slowly (Kamachi et al., 2001; Recio et al., 2013). Ekman (2003) suggested that, among all the expressions, sadness is the one lasting the longest over time, a property that may explain why slowness or stillness may increase recognition performance. Our results further confirm that the idiosyncratic properties of this expression are inherently slow.

We should acknowledge that we did not assess whether elderly people's cognitive abilities influenced recognition performance. Fluid intelligence (e.g., Horning et al., 2012; Sullivan & Ruffman, 2004a), processing speed (e.g., Orgeta & Phillips, 2007), verbal memory (e.g., MacPherson et al., 2002), or discrimination of visual information (Mill, Allik, Realo, & Valk, 2009) are all cognitive faculties that are critical for the recognition of human facial expressions and have been found to decrease with increasing age (e.g., Mill et al., 2009; Salthouse, 2004). Because our stimuli were only presented for 1 s, reduced processing speed in elderly people may have influenced their recognition performance for static face images as the number of facial features extracted from the faces in this limited presentation time might have been lower than that in younger adults. Interestingly, a recent cross-sectional study that examined the influence of different cognitive abilities on facial-expression recognition observed that these faculties contributed to the performance but did not fully account for the impairments observed in older adults (Horning et al., 2012). Additionally, in a later study, Zhao et al. (2016) observed that the slower processing speed in elderly people was not responsible for facial expression-recognition deficits (but see Suzuki & Akiyama, 2013; West et al., 2012).

It is also worth noting that differences in recognition abilities could stem from differences in the cohorts that

were tested, such as educational shifts, cultural norms, or social differences. To the best of our knowledge, no prior research has ever examined the extent to which the recognition of facial expression is influenced by these cohort effects (Ruffman et al., 2008).

The shuffled condition

We introduced this condition to fully replicate the study conducted by Gold et al. (2013) without having clear predictions for this experimental condition. This condition was originally designed to assess whether human observers were sensitive to the temporal development of an expression over time (i.e., order of frames). Our results revealed similar developmental trajectories for the six basic expressions in the shuffled condition (i.e., increase, peak efficiency, and decrease) although recognition rates were generally lower than those observed in the other conditions. More specifically, our findings revealed a recognition advantage for the dynamic expressions of anger, disgust, happiness, and surprise over the shuffled ones and a recognition advantage for the static expressions of happiness, disgust, and surprise over the shuffled ones. In contrast, we observed better recognition performance for the expression of fear in the shuffled compared to the static or dynamic conditions. As reported by our participants, this advantage for fear could be accounted for by the properties of the stimuli themselves. The shuffled expressions were generated by temporally randomizing the frames of the dynamic movies. This procedure leads to the impression that the actors performing the emotional expressions are shaking giving the feeling that they are afraid.

Interestingly, differences in recognition performance across conditions are inconsistent with the results reported by Gold et al. (2013), who observed similar performance in all three conditions. In that prior study, however, the authors did not consider the recognition rates of the individual expressions effectively, collapsing them across the six expressions in each condition. Our findings offer, therefore, new evidence that the temporal progression of information (i.e., the order of the frames) provided by genuine natural expressions is more important for the recognition of some expressions (e.g., anger, disgust, happiness, surprise) than others (e.g., fear, sadness). Given the very nonecological nature of the stimuli, we do not further discuss these results as their contribution is limited from a theoretical point of view.

Methodological considerations

In the current study, we used a hierarchical Bayesian model with weakly informative priors. The flexibility

and power of the Bayesian approach in dealing with time-series data and building nonlinear models was also recently demonstrated in emotion research (e.g., see Krone, Albers, Kuppens, & Timmerman, 2017, for an application in personal emotion dynamics) thanks to the rapid development in probabilistic language programming. It is worth noting that there are alternative candidate models that could capture the nonlinear relationship between age and some psychological or behavioral measurements, including latent growth curve models (for an introductory text, see Duncan, Duncan, & Strycker, 2013); generalized additive mixed models, including spline regression; and quadratic linear mixed-effect models (S. N. Wood, 2006). Some of these models have been previously applied to investigate similar questions, such as the estimation of the peak efficiency of diverse recognition abilities (i.e., change-point estimation, e.g., Cohen, 2012; Cudeck & Klebe, 2002). The Bayesian modeling framework we used here provides a coherent mathematical language to describe our model and assumption while giving the flexibility to potentially extend part of the component to build more complex models. Moreover, it allowed us to properly quantify the uncertainty and regulate the estimation across different conditions using hyper-priors.

To identify inverse U-shape patterns, such as those observed in our study, previous modeling methods occasionally involved the testing of a quadratic relationship (e.g., a significant regression coefficient of age²) even if such practice is not always valid (Simonsohn, 2017). Instead, Simonsohn (2017) suggested fitting two separate linear models and comparing the coefficients of the two slopes as a more valid alternative. Although our model is conceptually similar to Simonsohn's model, there are two major differences. First, the inference proposed by Simonsohn involved multiple model-fitting steps by initially identifying the break point (i.e., peak efficiency in our case) and then estimating the coefficients of the two linear functions. In contrast, with a full model that jointly estimates the break point and the linear functions, we could better estimate the parameters and quantify the associated uncertainty. Second, the intercepts of our step-linear function are linked and represented as one value (i.e., the recognition ability at peak efficiency), whereas in Simonsohn's model the two linear functions are not linked. The linked linear function is more appropriate in our case as it is unlikely to have a sudden increase or decrease in recognition ability in a short span during natural development. Nonetheless, an implicit yet important assumption present in both models is that the peak efficiency is found somewhere in the middle of the life span (or, more precisely, not at either of the two extrema). Indeed, if the peak efficiency is at the lower

or upper limit (e.g., too young or too old), the parameter estimation may not be accurate.

Our model estimation performed well except for the expression of happiness because of the ceiling effect we observed for this expression. The divergence in the trace and the multimodal in the posterior distribution of the peak efficiency both indicate that the current step-linear model is not the best suited to represent changes in recognition abilities across the life span for this expression. Currently, all the expressions are estimated independently. Although modeling this way is easier to interpret, we ignored the random effect in the subjects across the expressions. Future studies are necessary to take into account the random effect from each subject (intercept and slope). This could be done by directly modeling the full confusion matrix from each subject (instead of only looking at the diagonal in the current study), presumably with some matrix decomposition trick or a dirichlet-categorical model.

Finally, our model allowed us to estimate the overall advantage of one condition over another before and after the peak efficiency. However, because we decided to consider the age as a continuum and not rely on specific age groups on the basis of arbitrary boundaries, our model did not allow us to finely estimate at which precise age the dynamic advantage emerges or disappears.

Conclusions

Current knowledge about facial expression recognition primarily arises from studies that use static images. In our daily life, however, natural faces are dynamic; they evolve over time in some particular ways to convey crucial information for adapted social behaviors. Prior studies investigating the importance of dynamic cues for the processing of facial expressions have yielded equivocal results with some studies suggesting that dynamic expressions are more readily recognizable than static images and others suggesting that they are not. In order to clarify these results and to determine if age is a critical factor to account for such discrepancies, we conducted a large cross-sectional study to investigate the recognition of facial expressions by participants aged 5 to 96. More than 400 observers were instructed to categorize static, dynamic, and shuffled expressions according to the six basic expressions. Our findings revealed that, regardless of the age of the observer or temporal condition, happiness was the best recognized facial expression, whereas fear was the most difficult to effectively categorize as this expression was commonly confused with surprise. Bayesian modeling allowed us to quantify the steepness of increase and decrease in

performance for each individual expression in each condition. Our results also reveal a data-driven estimation of the peak efficiency for every expression and, finally, provide new evidence for a dynamic advantage for facial expression recognition, stronger for some expressions than others and more important around specific points in the life course. Notably, performance for static images was less effective in the elderly population. Altogether, our findings highlight the importance of using ecologically valid faces in exploring the recognition of facial expressions and invite caution while drawing conclusions from studies that use only static images to this aim.

Keywords: facial expressions of emotion, life span, static, dynamic, aging

Acknowledgments

We would like to thank Prof. Jason Gold and his colleagues from Indiana University, Bloomington, for providing us with the stimuli used in this study. Very special thanks go to all our participants, especially the children and adolescents who participated in our study from the following schools: Ecole Enfantine et Primaire de Marly Grand Pré and Cycle d'Orientation de Domdidier in the area of Fribourg, Switzerland. We would also like to thank all the teachers for their patience and help, especially the head teachers, Claude Meuwly from Marly and Chantal Vienny Guerry from Domdidier, as well as technician Zvonko Traykoski for meticulously organizing our visits and ensuring that everything worked well. We would also like to express our gratitude to the retirement homes Foyer de Bouleyres and Maison Bourgeoisiale in Bulle and Foyer St-Martin in Cottens, especially to Christian Rime, Véronique Castella, Isabelle Montagnon, and Philippe Bourquin for letting us test their residents. We would also like to thank Maria Teresa Turano and Prof. Maria Pia Viggiano from Florence, Léa Poitrine, Claudia Wyler, Qendresa Shkodra, Vanessa Ferrari, Pauline Rotztetter, Pauline Schaller, Linda Pigozzi, Lauriane Beffa, Hugo Najberg, Christel Aichele, and Martina Studer for their precious help with testing. This study was supported by grant F14/06 from the Rectors' Conference of Swiss Universities (CRUS) awarded to ARR and by a grant from the Swiss National Science Foundation (n°100014_156490/1) awarded to RC. The publication fees were covered by the Comité des Alumni et Amis de l'Université de Fribourg.

Commercial relationships: none.

Corresponding authors: Anne-Raphaëlle Richoz; Roberto Caldara.

Email: anne-raphaëlle.richo@unifr.ch; roberto.caldara@unifr.ch.
Address: Department of Psychology, University of Fribourg, Fribourg, Switzerland.

References

- Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005, January 6). A mechanism for impaired fear recognition after amygdala damage. *Nature*, *433*(7021), 68–72.
- Adolphs, R., Tranel, D., & Damasio, A. R. (2003). Dissociable neural systems for recognizing emotions. *Brain and Cognition*, *52*(1), 61–69.
- Alves, N. T. (2013). Recognition of static and dynamic facial expressions: A study review. *Estudos de Psicologia (Natal)*, *18*(1), 125–130.
- Ambadar, Z., Schooler, J. W., & Cohn, J. F. (2005). Deciphering the enigmatic face: The importance of facial dynamics in interpreting subtle facial expressions. *Psychological Science*, *16*(5), 403–410.
- Andersen, S. L. (2003). Trajectories of brain development: Point of vulnerability or window of opportunity? *Neuroscience & Biobehavioral Reviews*, *27*(1), 3–18.
- Arsalidou, M., Morris, D., & Taylor, M. J. (2011). Converging evidence for the advantage of dynamic facial expressions. *Brain Topography*, *24*(2), 149–163.
- Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., ... Bentin, S. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science*, *19*(7), 724–732.
- Back, E., Ropar, D., & Mitchell, P. (2007). Do the eyes have it? Inferring mental states from animated faces in autism. *Child Development*, *78*(2), 397–411.
- Barrett, L. F., & Kensinger, E. A. (2010). Context is routinely encoded during emotion perception. *Psychological Science*, *21*(4), 595–599.
- 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*(1862), 1–9.
- Bernstein, M., Erez, Y., Blank, I., & Yovel, G. (2017). The processing of dynamic faces in the human brain: Support for an integrated neural framework of face processing. *BioRxiv* :140939.
- Blakemore, S.-J. (2012). Imaging brain development: The adolescent brain. *NeuroImage*, *61*(2), 397–406.
- Blakemore, S.-J., & Choudhury, S. (2006). Development of the adolescent brain: Implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry*, *47*(3–4), 296–312.
- Bould, E., & Morris, N. (2008). Role of motion signals in recognizing subtle facial expressions of emotion. *British Journal of Psychology*, *99*(2), 167–189.
- Bould, E., Morris, N., & Wink, B. (2008). Recognising subtle emotional expressions: The role of facial movements. *Cognition and Emotion*, *22*(8), 1569–1587.
- Brainard, D. H., (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.
- Caldara, R. (2017). Culture reveals a flexible system for face processing. *Current Directions in Psychological Science*, *26*(3), 249–255.
- Calder, A. J., Keane, J., Manly, T., Sprengelmeyer, R., Scott, S., Nimmo-Smith, I., & Young, A. W. (2003). Facial expression recognition across the adult life span. *Neuropsychologia*, *41*(2), 195–202.
- Carlson, J. M., & Mujica-Parodi, L. R. (2015). Facilitated attentional orienting and delayed disengagement to conscious and nonconscious fearful faces. *Journal of Nonverbal Behavior*, *39*(1), 69–77.
- Casey, B., Tottenham, N., Liston, C., & Durston, S. (2005). Imaging the developing brain: What have we learned about cognitive development? *Trends in Cognitive Sciences*, *9*(3), 104–110.
- Chen, J. J., Rosas, H. D., & Salat, D. H. (2011). Age-associated reductions in cerebral blood flow are independent from regional atrophy. *NeuroImage*, *55*(2), 468–478.
- Christie, F., & Bruce, V. (1998). The role of dynamic information in the recognition of unfamiliar faces. *Memory & Cognition*, *26*(4), 780–790.
- Cohen, P. (2012). *Applied data analytic techniques for turning points research*. New York: Routledge.
- Cudeck, R., & Klebe, K. J. (2002). Multiphase mixed-effects models for repeated measures data. *Psychological Methods*, *7*(1), 41–63.
- Cunningham, D. W., & Wallraven, C. (2009). Dynamic information for the recognition of conversational expressions. *Journal of Vision*, *9*(13):7, 1–17, <https://doi.org/10.1167/9.13.7>. [PubMed] [Article]
- Delis, I., Chen, C., Jack, R. E., Garrod, O. G., Panzeri, S., & Schyns, P. G. (2016). Space-by-time manifold representation of dynamic facial expressions for emotion categorization. *Journal of Vision*, *16*(8):14, 1–20, <https://doi.org/10.1167/16.8.14>. [PubMed] [Article]
- De Sonneville, L., Verschoor, C., Njokiktjen, C., Op het Veld, V., Toorenaar, N., & Vranken, M. (2002).

- Facial identity and facial emotions: Speed, accuracy, and processing strategies in children and adults. *Journal of Clinical and Experimental Neuropsychology*, 24(2), 200–213.
- De Winter, F.-L., Van den Stock, J., de Gelder, B., Peeters, R., Jastorff, J., Sunaert, S., & hellip; Vandenbulcke, M. (2016). Amygdala atrophy affects emotion-related activity in face-responsive regions in frontotemporal degeneration. *Cortex*, 82, 179–191.
- Duncan, T. E., Duncan, S., & Strycker, L. A. (2013). *An introduction to latent variable growth curve modeling: Concepts, issues, and application*. New York: Routledge Academic.
- 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.
- Durston, S., Pol, H. E. H., Casey, B., Giedd, J. N., Buitelaar, J. K., & Van Engeland, H. (2001). Anatomical MRI of the developing human brain: What have we learned? *Journal of the American Academy of Child & Adolescent Psychiatry*, 40(9), 1012–1020.
- Ehrlich, S. M., Schiano, D. J., & Sheridan, K. (2000). Communicating facial affect: It's not the realism, it's the motion. In *Proceedings of ACM CHI 2000 Conference on Human Factors in Computing Systems* (pp. 252–253), New York, NY: ACM.
- Ekman, P. (2003). *Emotions revealed. Understanding faces and feelings*. London: The Orion Publishing Group Ltd.
- Ekman, P., & Friesen, W. V. (1976). Measuring facial movement. *Environmental Psychology and Nonverbal Behavior*, 1(1), 56–75.
- Fiorentini, C., & Viviani, P. (2011). Is there a dynamic advantage for facial expressions? *Journal of Vision*, 11(3):17, 1–15, <https://doi.org/10.1167/11.3.17>. [PubMed] [Article]
- Fiset, D., Blais, C., Royer, J., Richoz, A.-R., Dugas, G., & Caldara, R. (2017). Mapping the impairment in decoding static facial expressions of emotion in prosopagnosia. *Social Cognitive and Affective Neuroscience*, 12(8), 1334–1341.
- Folstein, M. F. (1975). A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189–198.
- Fox, C. J., Iaria, G., & Barton, J. J. (2009). Defining the face processing network: Optimization of the functional localizer in fMRI. *Human Brain Mapping*, 30(5), 1637–1651.
- Ganel, T., Valyear, K. F., Goshen-Gottstein, Y., & Goodale, M. A. (2005). The involvement of the “fusiform face area” in processing facial expression. *Neuropsychologia*, 43(11), 1645–1654.
- 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.
- Gao, X., & Maurer, D. (2010). A happy story: Developmental changes in children's sensitivity to facial expressions of varying intensities. *Journal of Experimental Child Psychology*, 107(2), 67–86.
- Gao, X., & Maurer, D. (2011). A comparison of spatial frequency tuning for the recognition of facial identity and facial expressions in adults and children. *Vision Research*, 51(5), 508–519.
- 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(14), R663–R664.
- Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4), 457–472.
- Gepner, B., Deruelle, C., & Grynfeldt, S. (2001). Motion and emotion: A novel approach to the study of face processing by young autistic children. *Journal of Autism and Developmental Disorders*, 31(1), 37–45.
- Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience*, 11(5), 473–490.
- 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(5):23, 1–12, <https://doi.org/10.1167/13.5.23>. [PubMed] [Article]
- Grainger, S. A., Henry, J. D., Phillips, L. H., Vanman, E. J., & Allen, R. (2015). Age deficits in facial affect recognition: The influence of dynamic cues. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 72(4), 622–632.
- Gross, A. L., & Ballif, B. (1991). Children's understanding of emotion from facial expressions and situations: A review. *Developmental Review*, 11(4), 368–398.
- Harwood, N. K., Hall, L. J., & Shinkfield, A. J. (1999). Recognition of facial emotional expressions from moving and static displays by individuals with mental retardation. *American Journal on Mental Retardation*, 104(3), 270–278.

- Herba, C., Landau, S., Russell, T., Ecker, C., & Phillips, M. L. (2006). The development of emotion-processing in children: Effects of age, emotion, and intensity. *Journal of Child Psychology and Psychiatry*, 47(11), 1098–1106.
- 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.
- Hof, P. R., & Morrison, J. H. (2004). The aging brain: Morphomolecular senescence of cortical circuits. *Trends in Neurosciences*, 27(10), 607–613.
- Horning, S. M., Cornwell, R. E., & Davis, H. P. (2012). The recognition of facial expressions: An investigation of the influence of age and cognition. *Aging, Neuropsychology, and Cognition*, 19(6), 657–676.
- Humphreys, G. W., Donnelly, N., & Riddoch, M. J. (1993). Expression is computed separately from facial identity, and it is computed separately for moving and static faces: Neuropsychological evidence. *Neuropsychologia*, 31(2), 173–181.
- Jack, C. R., Petersen, R. C., Xu, Y. C., Waring, S. C., O'Brien, P. C., Tangalos, E. G., ... Kokmen, E. (1997). Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease. *Neurology*, 49(3), 786–794.
- Jack, R. E., Garrod, O. G., & Schyns, P. G. (2014). Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time. *Current Biology*, 24(2), 187–192.
- Jiang, Z., Li, W., Recio, G., Liu, Y., Luo, W., Zhang, D., & Sun, D. (2014). Time pressure inhibits dynamic advantage in the classification of facial expressions of emotion. *PLoS One*, 9(6), 1–7.
- Johnston, P., Mayes, A., Hughes, M., & Young, A. W. (2013). Brain networks subserving the evaluation of static and dynamic facial expressions. *Cortex*, 49(9), 2462–2472.
- Kamachi, M., Bruce, V., Mukaida, S., Gyoba, J., Yoshikawa, S., & Akamatsu, S. (2001). Dynamic properties influence the perception of facial expressions. *Perception*, 30(7), 875–887.
- Kätsyri, J. (2006). *Human recognition of basic emotions from posed and animated dynamic facial expressions*. Unpublished doctoral dissertation, Helsinki University of Technology. Retrieved from <https://aaltodoc.aalto.fi/bitstream/handle/123456789/2807/isbn951228538X.pdf?sequence=1&isAllowed=y>
- Kätsyri, J., Saalasti, S., Tiippana, K., von Wendt, L., & Sams, M. (2008). Impaired recognition of facial emotions from low-spatial frequencies in Asperger syndrome. *Neuropsychologia*, 46(7), 1888–1897.
- Kätsyri, J., & Sams, M. (2008). The effect of dynamics on identifying basic emotions from synthetic and natural faces. *International Journal of Human-Computer Studies*, 66(4), 233–242.
- Kessler, H., Doyen-Waldecker, C., Hofer, C., Hoffmann, H., Traue, H. C., & Abler, B. (2011). Neural correlates of the perception of dynamic versus static facial expressions of emotion. *Psychosocial Medicine*, 8, 1–8.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, 36(14), 1–16.
- Knappmeyer, B., Thornton, I. M., & Bühlhoff, H. H. (2003). The use of facial motion and facial form during the processing of identity. *Vision Research*, 43(18), 1921–1936.
- Krendl, A. C., & Ambady, N. (2010). Older adults' decoding of emotions: Role of dynamic versus static cues and age-related cognitive decline. *Psychology and Aging*, 25(4), 788–793.
- Krone, T., Albers, C. J., Kuppens, P., & Timmerman, M. E. (2018). A multivariate statistical model for emotion dynamics. *Emotion*, 18(5), 739–754.
- Krumhuber, E. G., Kappas, A., & Manstead, A. S. (2013). Effects of dynamic aspects of facial expressions: A review. *Emotion Review*, 5(1), 41–46.
- Kruschke, J. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan* (2nd ed.). New York: Academic Press.
- LaBar, K. S., Crupain, M. J., Voyvodic, J. T., & McCarthy, G. (2003). Dynamic perception of facial affect and identity in the human brain. *Cerebral Cortex*, 13(10), 1023–1033.
- Leitzke, B. T., & Pollak, S. D. (2016). Developmental changes in the primacy of facial cues for emotion recognition. *Developmental Psychology*, 52(4), 572–581.
- Liang, Y., Liu, B., Xu, J., Zhang, G., Li, X., Wang, P., & Wang, B. (2017). Decoding facial expressions based on face-selective and motion-sensitive areas. *Human Brain Mapping*, 38(6), 3113–3125.
- LoBue, V. (2010). What's so scary about needles and knives? Examining the role of experience in threat detection. *Cognition and Emotion*, 24(1), 180–187.
- MacPherson, S. E., Phillips, L. H., & Della Sala, S. (2002). Age, executive function and social decision making: A dorsolateral prefrontal theory of cognitive aging. *Psychology and Aging*, 17(4), 598–609.
- Malatesta, C. Z., Izard, C. E., Culver, C., & Nicolich, M. (1987). Emotion communication skills in young,

- middle-aged, and older women. *Psychology and Aging*, 2(2), 193–203.
- Mather, M., Canli, T., English, T., Whitfield, S., Wais, P., Ochsner, K., . . . Carstensen, L. L. (2004). Amygdala responses to emotionally valenced stimuli in older and younger adults. *Psychological Science*, 15(4), 259–263.
- McDowell, C. L., Harrison, D. W., & Demaree, H. A. (1994). Is right hemisphere decline in the perception of emotion a function of aging? *International Journal of Neuroscience*, 79(1–2), 1–11.
- Mill, A., Allik, J., Realo, A., & Valk, R. (2009). Age-related differences in emotion recognition ability: A cross-sectional study. *Emotion*, 9(5), 619–630.
- Mitchell, A. J. (2009). A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. *Journal of Psychiatric Research*, 43(4), 411–431.
- Montirosso, R., Peverelli, M., Frigerio, E., Crespi, M., & Borgatti, R. (2010). The development of dynamic facial expression recognition at different intensities in 4- to 18-year-olds. *Social Development*, 19(1), 71–92.
- Moreno, C., Borod, J. C., Welkowitz, J., & Alpert, M. (1993). The perception of facial emotion across the adult life span. *Developmental Neuropsychology*, 9(3–4), 305–314.
- Navarro, C., & Gonzalo, L. M. (1991). Cambios en el complejo amigdalino humano debidos a la edad [Changes in the human amygdaloid complex due to age]. *Revista de Medicina de la Universidad de Navarra*, 35, 7–12.
- Nelson, N. L., Hudspeth, K., & Russell, J. A. (2013). A story superiority effect for disgust, fear, embarrassment, and pride. *British Journal of Developmental Psychology*, 31(3), 334–348.
- Nelson, N. L., & Russell, J. A. (2011). Putting motion in emotion: Do dynamic presentations increase preschooler's recognition of emotion? *Cognitive Development*, 26(3), 248–259.
- Orgeta, V., & Phillips, L. H. (2007). Effects of age and emotional intensity on the recognition of facial emotion. *Experimental Aging Research*, 34(1), 63–79.
- Paulmann, S., Jessen, S., & Kotz, S. A. (2009). Investigating the multimodal nature of human communication: An ERP study. *Journal of Psychophysiology*, 23(2), 63–76.
- Phelps, E. A., Ling, S., & Carrasco, M. (2006). Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychological Science*, 17(4), 292–299.
- Pitcher, D., Dilks, D. D., Saxe, R. R., Triantafyllou, C., & Kanwisher, N. (2011). Differential selectivity for dynamic versus static information in face-selective cortical regions. *NeuroImage*, 56(4), 2356–2363.
- Pourtois, G., de Gelder, B., Bol, A., & Crommelinck, M. (2005). Perception of facial expressions and voices and of their combination in the human brain. *Cortex*, 41(1), 49–59.
- Raz, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 1–90). Mahwah, NJ: Lawrence Erlbaum Associates.
- Recio, G., Schacht, A., & Sommer, W. (2013). Classification of dynamic facial expressions of emotion presented briefly. *Cognition & Emotion*, 27(8), 1486–1494.
- Richo, 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.
- Rodger, H., Lao, J., & Caldara, R. (2018). Quantifying facial expression signal and intensity use during development. *Journal of Experimental Child Psychology*, 174, 41–59.
- Rodger, H., Vizioli, L., Ouyang, X., & Caldara, R. (2015). Mapping the development of facial expression recognition. *Developmental Science*, 18(6), 926–939.
- Rossini, P. M., Rossi, S., Babiloni, C., & Polich, J. (2007). Clinical neurophysiology of aging brain: From normal aging to neurodegeneration. *Progress in Neurobiology*, 83(6), 375–400.
- Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. *Neuroscience & Biobehavioral Reviews*, 32(4), 863–881.
- Rymarczyk, K., Biele, C., Grabowska, A., & Majczynski, H. (2011). EMG activity in response to static and dynamic facial expressions. *International Journal of Psychophysiology*, 79(2), 330–333.
- Salthouse, T. A. (2004). What and when of cognitive aging. *Current Directions in Psychological Science*, 13(4), 140–144.
- Sato, W., Fujimura, T., & Suzuki, N. (2008). Enhanced facial EMG activity in response to dynamic facial

- expressions. *International Journal of Psychophysiology*, 70(1), 70–74.
- Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., & Matsumura, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: An fMRI study. *Brain Research Cognitive Brain Research*, 20(1), 81–91.
- Sato, W., & Yoshikawa, S. (2007). Spontaneous facial mimicry in response to dynamic facial expressions. *Cognition*, 104(1), 1–18.
- Schultz, J., & Pilz, K. S. (2009). Natural facial motion enhances cortical responses to faces. *Experimental Brain Research*, 194(3), 465–475.
- Simonsohn, U. (2017). Two-Lines: A valid Alternative to the Invalid Testing of U-Shaped Relationships with Quadratic Regressions. *SSRN Electronic Journal*, 1–29. Retrieved from: <https://ssrn.com/abstract=3021690>.
- Smith, F. W., & Rossit, S. (2018). Identifying and detecting facial expressions of emotion in peripheral vision. *PLoS One*, 13(5), e0197160.
- Sullivan, S., & Ruffman, T. (2004a). Emotion recognition deficits in the elderly. *International Journal of Neuroscience*, 114(3), 403–432.
- Sullivan, S., & Ruffman, T. (2004b). Social understanding: How does it fare with advancing years? *British Journal of Psychology*, 95(1), 1–18.
- Suzuki, A., & Akiyama, H. (2013). Cognitive aging explains age-related differences in face-based recognition of basic emotions except for anger and disgust. *Aging, Neuropsychology, and Cognition*, 20(3), 253–270.
- Sweeny, T. D., Suzuki, S., Grabowecky, M., & Paller, K. A. (2013). Detecting and categorizing fleeting emotions in faces. *Emotion*, 13(1), 76–91.
- Sze, J. A., Goodkind, M. S., Gyurak, A., & Levenson, R. W. (2012). Aging and emotion recognition: Not just a losing matter. *Psychology and Aging*, 27(4), 940–950.
- Tardif, C., Lainé, F., Rodriguez, M., & Gepner, B. (2007). Slowing down presentation of facial movements and vocal sounds enhances facial expression recognition and induces facial–vocal imitation in children with autism. *Journal of Autism and Developmental Disorders*, 37(8), 1469–1484.
- Thomas, L. A., De Bellis, M. D., Graham, R., & LaBar, K. S. (2007). Development of emotional facial recognition in late childhood and adolescence. *Developmental Science*, 10(5), 547–558.
- Trautmann, S. A., Fehr, T., & Herrmann, M. (2009). Emotions in motion: Dynamic compared to static facial expressions of disgust and happiness reveal more widespread emotion-specific activations. *Brain Research*, 1284, 100–115.
- Turano, M. T., Lao, J., Richoz, A. R., De Lissa, P., Desgosciu, S. B. A., Viggiano, M. P., & Caldara, R. (2017). Fear boosts the early neural coding of faces. *Social, Cognitive, and Affective Neuroscience*, 12(12), 1959–1971.
- Uono, S., Sato, W., & Toichi, M. (2010). Brief report: Representational momentum for dynamic facial expressions in pervasive developmental disorder. *Journal of Autism and Developmental Disorders*, 40(3), 371–377.
- Vicari, S., Reilly, J. S., Pasqualetti, P., Vizzotto, A., & Caltagirone, C. (2000). Recognition of facial expressions of emotions in school-age children: The intersection of perceptual and semantic categories. *Acta Paediatrica*, 89(7), 836–845.
- Wallraven, C., Breidt, M., Cunningham, D. W., & Bülthoff, H. H. (2008). Evaluating the perceptual realism of animated facial expressions. *ACM Transactions on Applied Perception*, 4(4), 1–20.
- Wehrle, T., Kaiser, S., Schmidt, S., & Scherer, K. R. (2000). Studying the dynamics of emotional expression using synthesized facial muscle movements. *Journal of Personality and Social Psychology*, 78(1), 105–119.
- West, J. T., Horning, S. M., Klebe, K. J., Foster, S. M., Cornwell, R. E., Perrett, D., Burt, D. M., & Davis, H. P. (2012). Age effects on emotion recognition in facial displays: From 20 to 89 years of age. *Experimental Aging Research*, 38(2), 146–168.
- Widen, S. C. (2013). Children’s interpretation of facial expressions: The long path from valence-based to specific discrete categories. *Emotion Review*, 5(1), 72–77.
- Widen, S. C., & Russell, J. A. (2015). Do dynamic facial expressions convey emotions to children better than do static ones? *Journal of Cognition and Development*, 16(5), 802–811.
- 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(3), 671–684.
- Williams, L. M., Mathersul, D., Palmer, D. M., Gur, R. C., Gur, R. E., & Gordon, E. (2009). Explicit identification and implicit recognition of facial emotions: I. Age effects in males and females across 10 decades. *Journal of Clinical and Experimental Neuropsychology*, 31(3), 257–277.
- Wood, A., Lupyan, G., Sherrin, S., & Niedenthal, P. (2016). Altering sensorimotor feedback disrupts visual discrimination of facial expressions. *Psychonomic Bulletin & Review*, 23(4), 1150–1156.

- Wood, A., Rychlowska, M., Korb, S., & Niedenthal, P. (2016). Fashioning the face: Sensorimotor simulation contributes to facial expression recognition. *Trends in Cognitive Sciences*, 20(3), 227–240.
- Wood, S. N. (2006). *Generalized additive models: An introduction with R*. Boca Raton, FL: CRC Press.
- Xu, X., & Biederman, I. (2010). Loci of the release from fMRI adaptation for changes in facial expression, identity, and viewpoint. *Journal of Vision*, 10(14):36, 1–13, <https://doi.org/10.1167/10.14.36>. [PubMed] [Article]
- Yang, T. T., Menon, V., Eliez, S., Blasey, C., White, C. D., Reid, A. J., . . . Reiss, A. L. (2002). Amygdalar activation associated with positive and negative facial expressions. *Neuroreport*, 13(14), 1737–1741.
- Zhao, M.-F., Zimmer, H. D., Shen, X., Chen, W., & Fu, X. (2016). Exploring the cognitive processes causing the age-related categorization deficit in the recognition of facial expressions. *Experimental Aging Research*, 42(4), 348–364.