

The use of visual information in the recognition of posed and spontaneous facial expressions

Camille Saumure

Department of Psychoeducation and Psychology,
Université du Québec en Outaouais,
Gatineau, Québec, Canada

Marie-Pier Plouffe-Demers

Department of Psychoeducation and Psychology,
Université du Québec en Outaouais,
Gatineau, Québec, Canada

Amanda Estéphan

Department of Psychoeducation and Psychology,
Université du Québec en Outaouais,
Gatineau, Québec, Canada

Daniel Fiset

Department of Psychoeducation and Psychology,
Université du Québec en Outaouais,
Gatineau, Québec, Canada

Caroline Blais

Department of Psychoeducation and Psychology,
Université du Québec en Outaouais,
Gatineau, Québec, Canada



Recognizing facial expressions is crucial for the success of social interactions, and the visual processes underlying this ability have been the subject of many studies in the field of face perception. Nevertheless, the stimuli used in the majority of these studies consist of facial expressions that were produced on request rather than spontaneously induced. In the present study, we directly compared the visual strategies underlying the recognition of posed and spontaneous expressions of happiness, disgust, surprise, and sadness. We used the Bubbles method with pictures of the same individuals spontaneously expressing an emotion or posing with an expression on request. Two key findings were obtained: Visual strategies were less systematic with spontaneous than with posed expressions, suggesting a higher heterogeneity in the useful facial cues across identities; and with spontaneous expressions, the relative reliance on the mouth and eyes areas was more evenly distributed, contrasting with the higher reliance on the mouth compared to the eyes area observed with posed expressions.

Introduction

The face is a powerful social medium of communication (Jack & Schyns, 2017). Notwithstanding the information it provides regarding the identity of an individual and their social status (e.g., gender, ethnicity, and age), the face also allows inferences about the mental states and emotions that others are experiencing (Adolphs, 2001; Krolak-Salmon, 2011). The recognition and interpretation of facial expressions play a major role in regulating social behavior (Chambon & Baudouin, 2009), and an alteration of these processes is often linked with deficits in social functioning (Hall et al., 2004; Hooker & Park, 2002). For instance, a clear deficit in facial-expression recognition has been described in individuals with schizophrenia (Chambon & Baudouin, 2009; Kring & Elis, 2013; Mandal, Pandey, & Prasad, 1998) and autism spectrum disorder (Harms, Martin, & Wallace, 2010), and linked with atypical visual-extraction strategies (Clark, Gosselin, & Goghari, 2013; Lee, Gosselin, Wynn, & Green, 2011; Spazio, Adolphs, Hurley, & Piven, 2007).

Citation: Saumure, C., Plouffe-Demers, M.-P., Estéphan, A., Fiset, D., & Blais, C. (2018). The use of visual information in the recognition of posed and spontaneous facial expressions. *Journal of Vision*, 18(9):21, 1–16, <https://doi.org/10.1167/18.9.21>.

<https://doi.org/10.1167/18.9.21>

Received November 1, 2017; published September 20, 2018

ISSN 1534-7362 Copyright 2018 The Authors



Several studies have examined the strategies developed by the visual system to perform this task (e.g., Blais, Fiset, Roy, Saumure, & Gosselin, 2017; Blais, Roy, Fiset, Arguin, & Gosselin, 2012; Dailey et al., 2010; Elfenbein, Beaupré, Lévesque, & Hess, 2007; Fiset et al., 2017; Smith, Cottrell, Gosselin, & Schyns, 2005; Smith, & Merlusca, 2014; Sullivan, Ruffman, & Hutton, 2007; Thibault, Levesque, Gosselin, & Hess, 2012). This research has mainly focused on the study of posed facial expressions—that is, expressions exhibited on request. This body of research has uncovered the use of specific visual features for the recognition of each basic facial expression: for instance, the eyes for fear (Adolphs et al., 2005; Smith et al., 2005), the mouth for happiness (Dunlap, 1927; Smith et al., 2005), and the eyebrows, forehead, and eyes for sadness (Eisenbarth & Alpers, 2011; Smith et al., 2005). The mouth has also been shown to be the most useful area for discriminating all the posed basic expressions from one another (Blais et al., 2012; Duncan et al., 2017). However, few studies have assessed how spontaneous expressions are actually decoded. Here we define spontaneous expressions as natural ones that are displayed by an individual without another person requesting such a display (for a similar definition, see Matsumoto, Ollide, Schug, Willingham, & Callan, 2009). Previous studies that investigated the decoding of these expressions mostly focused on verifying whether individuals agree on which label to assign to a specific spontaneous expression, and showing a lower level of agreement compared with posed expressions (for a review, see Kayyal & Russell, 2013).

In investigating the transmission instead of the decoding (the latter being the focus of the current work), some studies have verified which facial cues are typically observed for both posed and spontaneous facial expressions. They also have demonstrated several differences between these types of expressions. For instance, some have revealed differences in the dynamic unfolding of posed and spontaneous expressions, showing that muscle activity is first initiated on the left side of the face for spontaneous expressions and on the right side of the face for the posed ones (Ross, Prodan, & Monnot, 2007; Ross & Pulusu, 2013). In addition to the differences observed with dynamic expressions, some have been observed with static expressions. For instance, computational studies have investigated the visual information contained in expressive faces that can be used by artificial vision systems to discriminate posed and spontaneous facial expressions. Results suggest that certain subareas of the face—namely the left brow, left eye, mouth, and chin—were given more weight when the computer program was assigned to discriminate posed from spontaneous facial expressions (Gan, Wu, Wang, & Ji, 2015). These regions might therefore contain information that can be helpful for

this specific task. Moreover, posed expressions usually exaggerate features (Kayyal & Russell, 2013), and are therefore more intense than spontaneous expressions. Spontaneous expressions also often include muscle activity that is not related to the experienced emotion, in part because social norms may dictate which expression is or is not appropriate in a specific context (Ekman, 1972) or because a person may experience more than one emotion at a time. Muscle activity that is not related to the dominant experienced emotion may decrease the signal clarity in comparison with posed expressions (Matsumoto et al., 2009). Consequently, the decrease in signal intensity and clarity that occurs with spontaneous expressions may affect the perceptual strategies underlying the decoding of spontaneous static expressions.

Relatedly, it has been proposed that the mouth area is more likely to be modulated when an expression is voluntarily changed to conform to social norms (Ekman & Friesen, 2003). Thus, although the mouth area is the most useful in discriminating posed basic facial expressions, this may not be the case for spontaneous expressions. In fact, for spontaneous expressions the mouth area is more likely to transmit inaccurate information about the emotion felt by an individual. To the best of our knowledge, no study has verified what visual information individuals rely on to correctly attribute a label to a spontaneous expression. Therefore, the present study aims at filling that gap by comparing visual-information use during the categorization of spontaneous and posed facial expressions of emotions.

Experiment

Visual-information use during the categorization of spontaneous and posed facial expressions of emotions was measured using the Bubbles technique (Gosselin & Schyns, 2001). It consists of randomly sampling the visual information contained in a stimulus, in the present case a facial expression. On each trial, a random subset of the visual information is rendered available to the participant. Participants' performance with these subsets of information allows us to infer which part of a stimulus they are using to perform the task accurately.

Method

Participants

Twenty participants (all White; 18 female, two male; average age = 21.5 years) took part in this experiment. The number of participants was determined based on previous studies using a similar method (e.g., Lee et al.,

2011). Because the method is based on a random sampling of visual information, it requires a very high number of trials in order to increase the signal-to-noise ratio. The high number of trials is typically reached by testing either a high number of participants on few trials or, alternatively, a low number of participants on many trials. Researchers usually collect between 5,000 (e.g., Smith & Merlusca, 2014) and 16,000 trials (e.g., Smith et al., 2005) per emotion. In the present study, we collected 10,000 trials per emotion, which falls in the typical range. All participants had normal or corrected-to-normal visual acuity. The study was approved by the Université du Québec en Outaouais's Research Ethics Committee and was conducted in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

Materials and stimuli

Stimuli were displayed on a calibrated high-resolution LCD monitor with a refresh rate of 60 Hz. The experimental program was written in MATLAB (MathWorks, Natick, MA), using functions from the Psychophysics Toolbox (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). The stimuli face width subtended around 7° of visual angle at a viewing distance of approximately 41 cm.

Stimuli consisted of pictures of White faces drawn from the MUG face database (Aifanti, Papachristou, & Delopoulos, 2010). This database is composed of pictures of individuals ($N = 86$) posing with six basic emotions (anger, disgust, fear, happiness, sadness, and surprise) and image sequences of some of those individuals ($N = 82$) captured while they viewed movies created to induce those six basic emotions. Thus, this database offers the advantage that the spontaneous and posed expressions were produced by the same individuals. This facilitates the direct comparison of the visual strategies used with the two kinds of expressions; that is, if some differences are revealed, they would not be attributable to differences inherent to the sample of featured individuals.

All the pictures displaying posed expressions were first inspected to eliminate the ones in which the photographed individuals were wearing artifacts or facial hair (e.g., glasses, jewels, beard) that could not be easily removed using Photoshop. The video sequences representing the individuals that were not eliminated in that first screening were then scanned to eliminate the ones in which the filmed person was positioned such that the camera did not capture a good frontal view of them. The remaining sequences ($N = 50$) were then scanned frame by frame (on average, 1,402 frames per video) to find the ones that best represented each of the basic emotions. An effort was made to find individuals for whom all the emotions were expressed during the

video sequence, and to keep the same number of male and female faces in the final stimulus sample. The emotions of fear and anger were not expressed frequently enough across individuals, so they were eliminated. At the end, pictures representing faces of 21 individuals (10 women, 11 men) expressing spontaneous and posed happiness, sadness, disgust, and surprise were kept for the present study. All the pictures selected for the experiment were transformed into grayscale images with a homogeneous gray background. Their luminance was normalized using the SHINE toolbox (Willenbockel et al., 2010). The images were also spatially aligned on the positions of the main internal facial features (eyes, mouth, and nose) using translation, rotation, and scaling.

Preliminary analyses on the selected stimuli

In order to make sure that the pictures selected were recognizable but also elicited a level of endorsement comparable to the one usually found in studies using spontaneous expressions, a separate sample of 30 White participants was asked to rate the perceived intensity of happiness, sadness, disgust, and surprise for each of the pictures (without bubbles) used in the present study. More specifically, each of the 168 pictures (21 identities \times 4 emotions \times 2 types of expression) was presented one at a time on the left side of the computer screen, while four scales (one per emotion) ranging from 1 (*not visible at all*) to 7 (*extremely visible*) were presented on the right side of the screen. The level of endorsement of the predicted label was measured by calculating the proportion of expressions that were correctly labeled (with the highest rating corresponding to the label a priori attributed to that expression). On average, a higher level of endorsement was found with posed expressions ($M = 92.9\%$, $SD = 13.6\%$) than with spontaneous expressions ($M = 66.8\%$, $SD = 33.8\%$), $t(83) = -7.29$, $p < 0.001$, 95% confidence interval (CI) $[-33.29, -19.01]$. Studies that have previously assessed the recognizability of spontaneous expressions have revealed low levels of endorsement of predicted labels, ranging from 15% to 66% (for a review, see Kayyal & Russell, 2013). Thus, the pictures selected for the present study corresponded to the higher bound of the label-endorsement range observed in previous studies.

Using this same set of data, the levels of intensity and ambiguity of the posed and spontaneous stimuli were compared. In fact, as explained in the Introduction, these are two characteristics that may be expected to differ between posed and spontaneous expressions. For each of the 84 pictures (21 identities \times 4 expressions), the highest and second-highest ratings were compared for posed and spontaneous expressions. The results indicated that the dominant emotion in each picture was judged as more intense for posed ($M =$

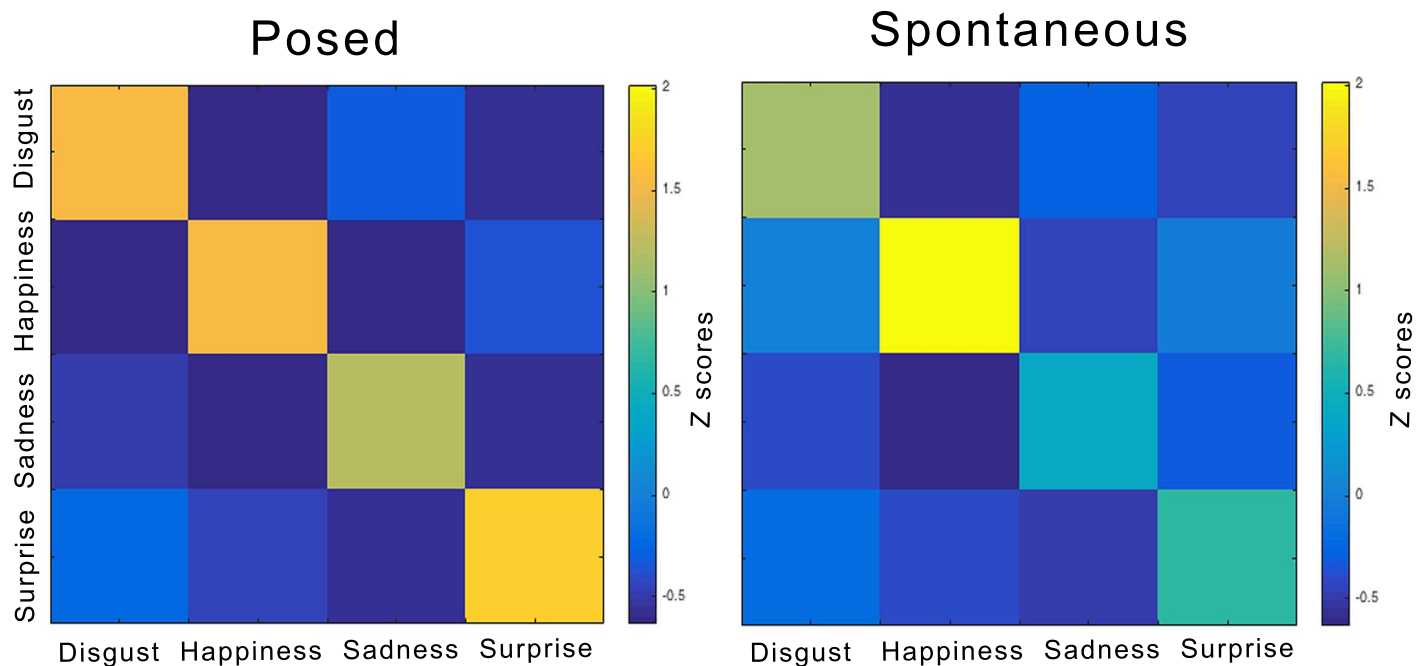


Figure 1. Ratings (in z scores) given on each emotion scale for the four posed and spontaneous expressions.

1.35, $SD = 0.21$) than for spontaneous ($M = 1.10$, $SD = 0.35$) expressions, $t(83) = 6.44$, $p < 0.001$, 95% CI [0.17, 0.34], but that the emotion perceived as the second most dominant was rated as being more intense with spontaneous ($M = -0.11$, $SD = 0.30$) than with posed ($M = -0.27$, $SD = 0.26$) expressions, $t(83) = -3.93$, $p < 0.001$, 95% CI [-0.24, -0.08]. Note that the values reported are in z scores; the raw scores were transformed into z scores to avoid potential individual rating biases and to highlight how the ratings differed from one another across the four scales. Figure 1 displays the ratings given to each emotion scale for the four expressions presented during the experiment. These results suggest not only that the dominant emotion in posed expressions was perceived as more intense but also that spontaneous emotions were more ambiguous than posed expressions; they contained a combination of facial cues from different emotions.

Bubbles method

The selected pictures were then used in the main experiment along with the Bubbles technique to reveal the visual information (in terms of spatial coordinates and spatial frequencies) useful for the categorization of posed and spontaneous facial expressions. On each trial, the creation of a “bubbled” stimulus went as follows: First, the image of a facial expression was decomposed into five spatial-frequency bands (2.2–4.5, 4.5–8.9, 8.9–17.9, 17.9–35.8, and 35.8–71.5 cycles/face; the remaining low-frequency bandwidth served as a constant background; see Figure 2, top row) using the

Laplacian pyramid (Burt & Adelson, 1983) included in the pyramid toolbox for MATLAB (Simoncelli, 1999). Then, independently for each spatial-frequency band, the bubbles’ locations (a bubble is a Gaussian aperture through which the information is visible) were randomly selected (see Figure 2, middle row). The size of the bubbles (full width at half maximum = 14.1, 28.3, 56.5, 113.0, and 226.1 pixels) was adjusted as a function of the frequency band so that each bubble revealed 1.5 cycles of spatial information. Because the size of the bubbles increased as the spatial scale became coarser, the number of bubbles differed across scales to keep a constant sampled area size across frequency bands. A point-wise multiplication was then performed between the bubble masks and the corresponding spatial-frequency bands of the filtered image (see Figure 2, bottom row). Finally, the information revealed by the bubbles was fused across the five frequency bands to produce an experimental stimulus (see Figure 2, bottom row, right image). Note that since the bubbles’ locations vary randomly across trials, it is possible after many trials to verify the statistical link between the visibility of a pixel or group of pixels and the probability of a correct answer.

Procedure

Before the experiment started, participants took part in a practice phase in which the unaltered faces (i.e., with no bubbles) were presented. Posed and spontaneous expressions were presented in separate blocks of 168 trials. This practice phase had two aims. First, it

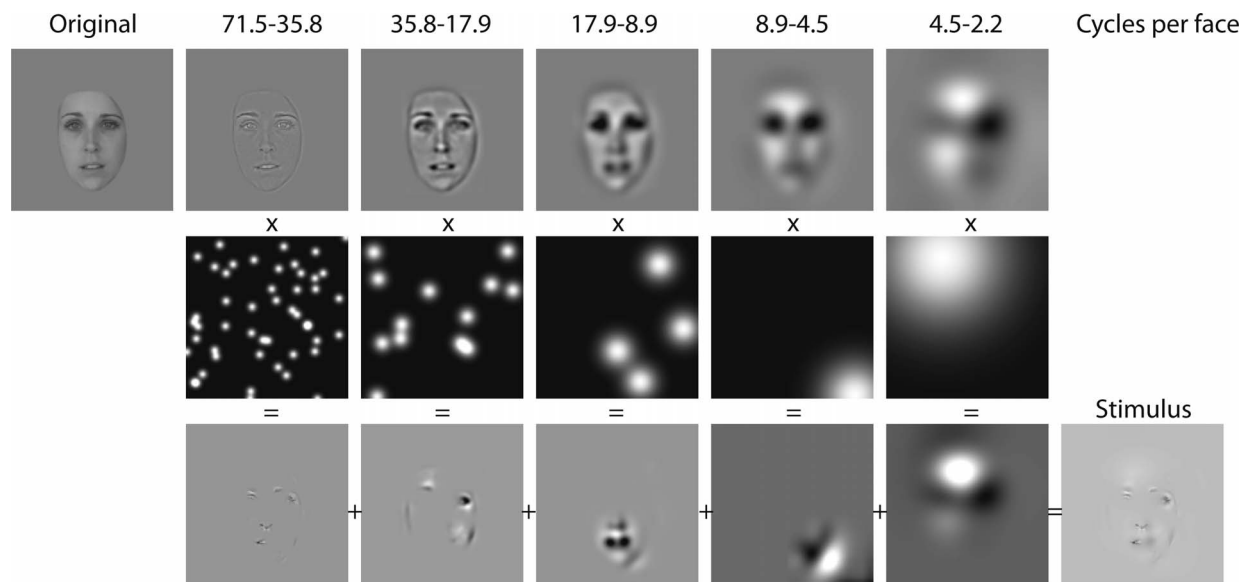


Figure 2. Illustration of the creation of a stimulus sampled with bubbles. See main text for more details. Note that the face stimuli used in this and all following figures are not part of the MUG database, for copyright reasons; they were instead taken from the first author, who gave her written consent for us to use her picture in this article.

allowed us to ensure that participants were able to easily recognize the expressions when they were presented without bubbles. In fact, the Bubbles technique assumes that the subset of visual information sampled on each trial modulates the probability of a correct recognition. For this reason, the unaltered stimuli must be recognizable by the participants, otherwise the probability of a correct answer would be modulated by two factors: the visual information available and the recognizability of the expression when no bubbles are applied to it. The unaltered stimuli must therefore attain a certain level of recognizability in order for us to be able to isolate the impact of the bubbles on performance. The second aim of the practice phase was to get participants to learn the keyboard keys associated with each of the four emotions. To complete the practice phase, participants needed to reach an accuracy criterion of 90% with both types of expressions.

In the Bubbles experiment, posed and spontaneous expressions were presented in separate blocks of 84 trials. Each participant completed 24 blocks with each type of expression, for a total of 4,032 trials. On each trial, the sequence of events unfolded as follows: First, a fixation cross appeared in the center of the computer screen for a duration of 500 ms. The fixation cross was quickly replaced by the stimulus (i.e., a bubbled face expressing one of the four emotions), which remained visible until the participant's response. The task was to categorize the expression by pressing the corresponding keyboard key. Blocks featuring posed and spontaneous expressions were alternated, and the order of the blocks was counterbalanced across participants. The number

of bubbles was adjusted separately for these two types of expressions using QUEST (Watson & Pelli, 1983) to maintain an average performance of 62.5% (i.e., halfway between chance and perfect performance) across the four expressions. No feedback was provided.

Results

During the practice phase, all participants reached the accuracy criterion of 90% within the first block with posed expressions, whereas they needed on average five blocks ($SD = 2.5$) to reach the same criterion with spontaneous expressions. During the experiment, means of 64.0 ($SD = 15.6$) and 34.4 ($SD = 8.7$) bubbles were necessary to maintain an average performance of 62.5% with spontaneous and posed stimuli, respectively. This indicates that more practice, $t(19) = 7.17$, $p < 0.001$, 95% CI [2.87, 5.23], and more visual information, $t(19) = 15.07$, $p < 0.001$, 95% CI [25.5, 33.7], were needed with spontaneous than with posed expressions to achieve a comparable accuracy rate. Indeed, it has recently been shown that the number of bubbles strongly correlates with the performance obtained with unaltered whole-face face stimuli (Royer, Blais, Goselin, Duncan, Fiset, 2015). Mean accuracy rates for each expression are reported in Table 1.

Computing the classification images

With the Bubbles technique, the utilization of visual information is quantified by computing classification images (CIs). CIs represent the weighted sums of all the

Expression	Posed	Spontaneous
Disgust	55.31% (7.07%)	55.97% (6.55%)
Happiness	69.40% (5.20%)	77.09% (6.30%)
Sadness	66.30% (6.13%)	54.06% (4.90%)
Surprise	63.39% (5.75%)	67.64% (5.81%)

Table 1. Mean (standard deviation) accuracy rate for each posed and spontaneous expression.

bubble masks presented during a given condition of an experiment (e.g., all the trials where a posed happy expression was presented to the participant), using as weights the participant's accuracy on each trial, transformed into z scores. Thus, computing a CI amounts to performing a multiple regression in which the bubble masks are the explanatory variable and the accuracy is the dependent variable.

First, a CI was computed for each participant, each facial emotion, and each expression type. Note that the CIs were summed across the five spatial-frequency bands (for the same procedure, see also Blais et al., 2012; Royer et al., 2016). Each CI was smoothed using a Gaussian kernel with a standard deviation of 12 pixels (full width at half maximum = 28.3 pixels). They were then transformed into z -score values using the mean and standard deviation of the null-hypothesis distribution, found by computing CIs with permuted z -score accuracy vectors. A 2 (types of expression: posed vs. spontaneous) \times 4 (emotions: disgust, happiness, surprise, sadness) repeated-measures analysis of variance (ANOVA) was conducted pixel by pixel to find the impact of the two factors and their interaction on information utilization. A statistical threshold was obtained using the Cluster test from the Stat4CI toolbox (Chauvin, Worsley, Schyns, Arguin, & Gosselin, 2005), a statistical method based on the random-fields theory that corrects for multiple comparisons (i.e., one ANOVA per pixel) by controlling for the family-wise error rate while taking into account the fact that contiguous pixels are not independent (i.e., they may be part of the same facial feature).

There were significant main effects of type of expression, $F_{\text{crit}}(1, 19) = 3.0$, $k = 4,748$, $p < 0.05$, and emotion, $F_{\text{crit}}(3, 57) = 3.0$, $k = 2,363$, $p < 0.05$. However, there was no significant interaction of the two factors for the utilization of facial information, $F_{\text{crit}}(3, 57) = 3.0$, $k = 2,363$. Figure 3 displays the maps of F scores for each effect. The areas for which a factor had a significant impact on information utilization are circled in white. Overall, both eyes, the folds between both eyebrows, and the mouth areas were more efficiently used with posed than with spontaneous expressions. The use of the cheeks and nasolabial folds varied as a function of the emotion presented.

The fact that the Expression type \times Emotion interaction was not significant suggests that the specific

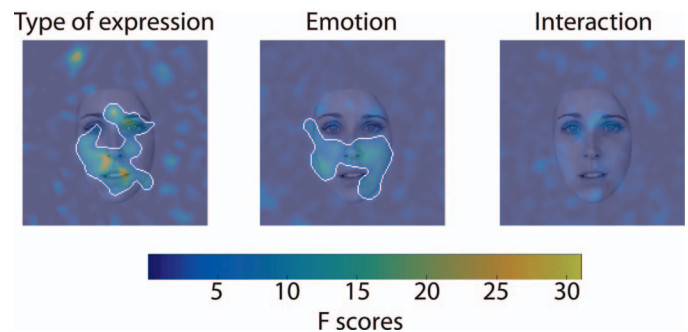


Figure 3. Maps of F scores indicating the impact of the type of expression, the emotion, and their interaction on the utilization of visual information during the recognition of facial expressions. Significant areas are circled in white.

pattern of visual information used with each emotion was similar for posed and spontaneous expressions. Nevertheless, to compare the present results with previous studies measuring the visual information used to recognize posed facial expressions (e.g., Blais et al., 2017; Smith et al., 2005; Smith & Merlusca, 2014), one-sample t tests were performed on each emotion of each expression type, to reveal what visual information significantly correlated with accuracy. Again, the significance threshold was obtained using a Cluster test, $t_{\text{crit}}(19) = 3.0$, $k = 620$, $p < 0.025$ (with a Bonferroni correction across the eight classification images). The results are displayed in Figure 4. The clusters significantly correlated with accuracy are circled in white. For the posed facial expressions, the results are consistent with previous studies: the mouth and nasolabial folds for disgust, and the mouth for happiness, sadness, and surprise. The eyes were also useful for posed disgust. Moreover, this strategy was relatively stable across participants. In fact, for each participant separately, we verified whether the pixels that were among the 10%

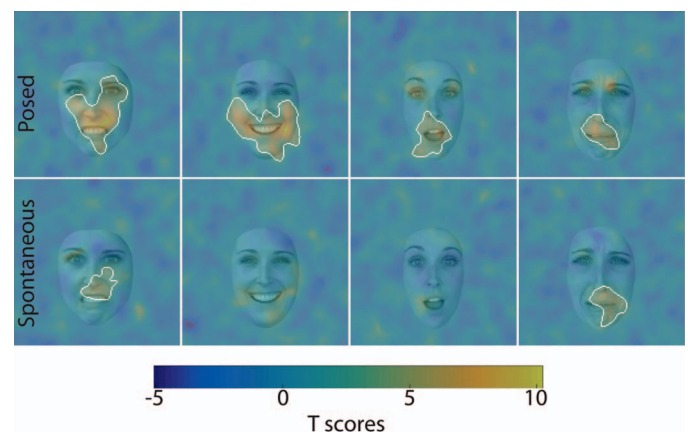


Figure 4. Maps indicating the relation between the utilization of visual information and the accuracy in recognizing each facial emotion, in their posed and spontaneous versions. Significant areas are circled in white.

with the highest z -score values overlapped with the significant area revealed for each expression; this was the case for 15, 16, 14, and 13 participants out of 20 for, respectively, the disgusted, happy, surprised, and sad expressions. Supplementary Figure S1 displays the homogeneity of the significantly useful areas across participants. Note that the eyes were also useful for posed disgust. Although this is not consistent with studies using the six basic emotion categories, it is consistent with studies in which anger was not part of the emotion sample (Smith & Merlusca, 2014), most likely because the eye area is not helpful in distinguishing between anger and disgust (e.g., Jack, Blais, Scheepers, Schyns, & Caldara, 2009).

For the spontaneous expressions, the mouth and nasolabial folds were significantly correlated with participants' accuracy at recognizing disgust, and the mouth area was correlated with their accuracy at recognizing sadness. This strategy was used by around half of the participants: 11 out of 20 for disgust and 10 out of 20 for sadness. No area reached significance for spontaneous surprise and happiness, although the area reaching the highest z -score values was the mouth, as was the case with posed expressions.

Relative utilization of the mouth and eye areas

The main effect of type of expression indicates that overall, most facial features were used more systematically to recognize posed than spontaneous expressions. However, because this result was obtained from a pixel-based analysis, it does not allow for comparison of the relative utilization, within one type of expression, of the different facial features. Since it has previously been observed that the mouth is the most diagnostic area to discriminate across all posed basic expressions, the relative utilization of the mouth and eye areas was compared here for both kinds of facial expressions.

In order to do so, CIs were computed in which the four expressions were pooled together (summed, and divided by the square root of the number of expressions; for a similar procedure, see Blais et al., 2012) before smoothing was applied. Smoothing was then applied using the same Gaussian kernel as in the preceding analysis. The CIs were finally transformed into z scores using the same procedure as in the preceding analysis—that is, using the mean and standard deviation of the null-hypothesis distribution, found by computing CIs with permuted z -score accuracy vectors. Then, a region-of-interest (ROI) analysis was conducted on these individual CIs. The analysis went as follows. For each participant, the maximum z -score value obtained in the mouth and eye areas of their CI was calculated. A 2 (ROI: mouth vs. eyes) \times 2 (type of expression: posed vs. spontaneous) repeated-measures ANOVA was conducted on these

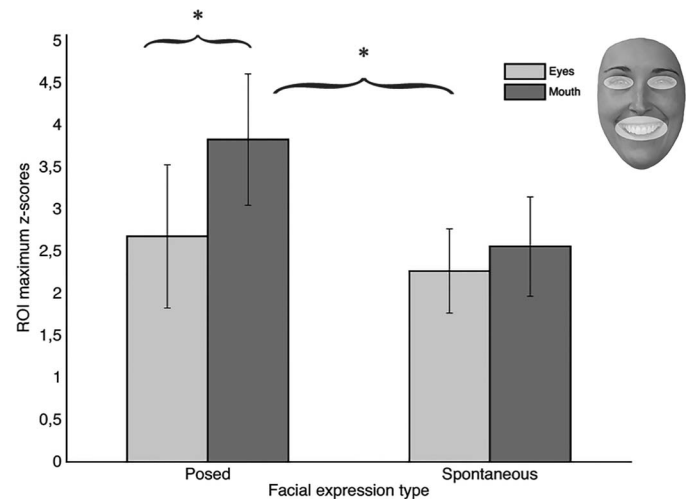


Figure 5. Average value of the maximum z scores reached in the mouth and eye areas. Error bars represent standard deviation. Inset represents the masks used for the region-of-interest analysis.

maximum z -score values. The results, as well as an illustration of the ROI used, are displayed in Figure 5. A significant ROI \times Type of expression was observed, $F(1, 19) = 7.87$, $p = 0.011$, so paired t tests were conducted. The results indicated that the mouth area was more useful than the eye area for posed facial expressions, $t(19) = -4.87$, $p < 0.001$, 95% CI $[-1.64, -0.66]$, but not for spontaneous ones, $t(19) = -1.8$, $p = 0.087$, 95% CI $[-0.66, 0.05]$. Moreover, while the mouth area was significantly more useful with posed than with spontaneous facial expressions, $t(19) = 6.2$, $p < 0.001$, 95% CI $[0.88, 1.65]$, the utilization of the eye area did not significantly differ between both types of expression, $t(19) = 1.8$, $p = 0.089$, 95% CI $[-0.07, -0.88]$.

Discussion

The pixel-based analyses indicated no interaction between the type of expression and the facial emotion processed. This suggests that the visual information processed for each emotion is similar for posed and spontaneous expressions. However, the main effect of emotion was significant, indicating that different facial cues were used to recognize each emotion. As already mentioned, the facial cues revealed for each posed emotion were consistent with previous studies. Finally, the main effect of the type of expression was significant, indicating that the z scores were overall lower for spontaneous than for posed expressions. Lower z scores suggest a less systematic visual strategy, which in turn may be the result of the higher ambiguity or lower intensity of spontaneous expressions.

Additionally, a comparison of the relative utilization of the mouth and eye areas indicated a decrease in the usefulness of the mouth area with spontaneous

compared with posed expressions. In fact, the maximum z -score values obtained in the eye area did not significantly differ from the mouth area for spontaneous expressions. This finding could indicate that the visual signal contained in this kind of expression changes in a way that modulates the relative informativeness of the mouth and eye areas compared with what is observed with posed expressions (see also Blais et al., 2012).

To verify whether the less systematic strategies observed with spontaneous expressions are indeed the result of a weaker signal (i.e., more ambiguous, less intense) and the relative informativeness of the mouth and eye areas changes with spontaneous expressions, we conducted a model-observer analysis using the same Bubbles task as the one performed by our participants.

Model observers

Participants' visual-information utilization, revealed with the Bubbles method, reflects an interaction between the diagnostic visual information for a task and the constraints imposed by the human visual system (Gosselin & Schyns, 2002). By contrast, the model-observer analysis that was conducted here imposed no visual constraint, which allowed us to dissociate the part of the visual strategies revealed with human participants that reflects the constraints of the visual system from the part that purely reflects the informativeness of the visual information. In other words, it allowed us to verify whether the lower z scores and the reduced reliance on the mouth area observed with spontaneous expressions could be predicted by the visual signal contained in this type of expression.

A model observer was tested with the same tasks as our 20 human participants and executed the same number of times as the number of human participants: Each instantiation of the model represented a unique participant. More specifically, to enable us to conduct the same analysis as with the human participants, for each model instantiation the same parameters were used that made up every trial for a given participant (i.e., bubbles mask, expression type, emotion category). Moreover, to further distinguish between the impact of stimulus ambiguity and intensity, the model observers were tested with an additional stimulus type: images of spontaneous emotions that were manipulated using the Abrasoft Fantamorph software to increase their intensity, in a linear fashion, to 150% of that of the original ones.

At the end, this provided us with the data of 20 different instantiations of the model that were tested with posed, spontaneous, and "high-intensity" spontaneous stimuli. The model observers' performance was maintained at the average accuracy rate obtained by the

participants with each expression by adding Gaussian white noise to the stimuli. The signal-to-noise ratio was manipulated on a trial-by-trial basis using QUEST.

The experiment went as follows. On each trial, a stimulus was created using the same parameters as in the main experiment. For instance, the emotion and bubbles mask used on Trial 1 for Participant 1 were used on Trial 1 for Model Observer 1, and so on for all trials of all participants. For the high-intensity spontaneous stimuli, the same parameters were used as with the spontaneous stimuli. Gaussian white noise was added to the stimulus before it was bubbled. All the face stimuli of the same expression type (e.g., spontaneous if a spontaneous expression was selected as the target) were then bubbled using the same bubbles mask as the target face. A pixel-by-pixel correlation between the bubbled (and noisy) target face and all the bubbled faces was performed. The emotion of the face stimulus that reached the highest correlation with the target face was considered the model observer's response (for the same kind of model observer, see Blais et al., 2012; Smith et al., 2005).

Computing the CIs

CIs were produced using the same procedure as the one described for the human participants. First, an analysis was conducted to compare the facial information used by the model observers with posed and spontaneous expressions. A 2 (types of expression: posed vs. spontaneous) \times 4 (emotions: disgust, happiness, surprise, sadness) repeated-measures ANOVA was conducted pixel by pixel to find the impact of the two factors and their interaction on the information utilization. A statistical threshold was obtained using the Cluster test from the Stat4CI toolbox. The main effect of type of expression was significant, $F_{\text{crit}}(1, 19) = 3.0$, $k = 3,039$, $p < 0.05$. However, there was no effect of emotion, and no significant interaction of the two factors on utilization of facial information, $F_{\text{crit}}(3, 57) = 3.0$, $k = 1,550$.

Two more analyses were performed to compare the visual information used by model observers with posed and high-intensity spontaneous expressions as well as with the spontaneous and high-intensity spontaneous expressions. Two 2 (types of expression) \times 4 (emotions: disgust, happiness, surprise, sadness) repeated-measures ANOVAs were conducted pixel by pixel to find the impact of the two factors and their interaction on information utilization. Statistical thresholds were obtained using the Cluster test from the Stat4CI toolbox. The results were highly consistent with those obtained in the preceding analysis. In comparing the posed and high-intensity spontaneous expressions, the main effect of type of expression was significant, $F_{\text{crit}}(1,$

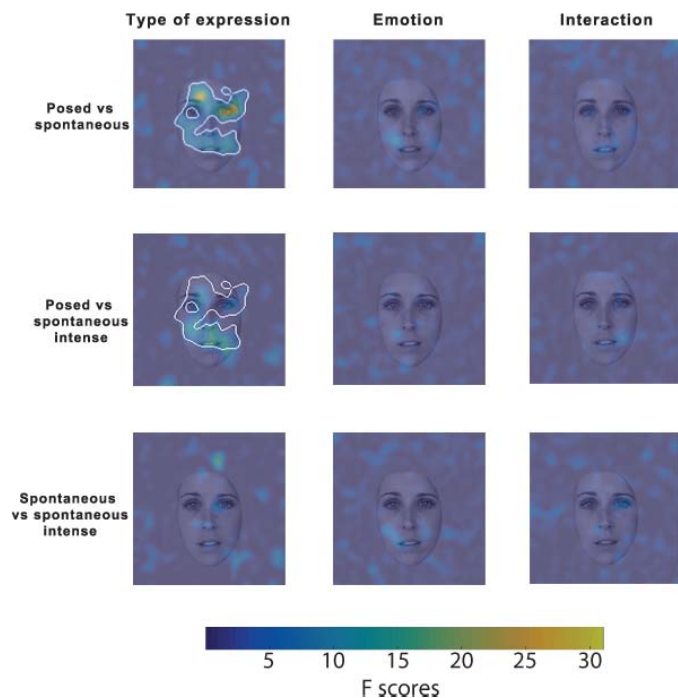


Figure 6. Maps of F scores indicating the impact of the type of expression, the emotion, and their interaction on the model observers' utilization of visual information during the recognition of facial expressions. Significant areas are circled in white.

19) = 3.0, $k = 3,039$, $p < 0.05$, but neither the effect of emotion nor the interaction of the two factors had a significant impact on the utilization of facial information, $F_{crit}(3, 57) = 3.0$, $k = 1,550$. However, comparing spontaneous with high-intensity spontaneous expressions, significance was not reached by the main effect of type of expression, $F_{crit}(1, 19) = 3.0$, $k = 3,039$; the main effect of emotion; or the interaction between both factors, $F_{crit}(3, 57) = 3.0$, $k = 1,550$.

Figure 6 displays the maps of F values for each effect in each ANOVA. The areas for which a factor had a significant effect on utilization are circled in white. Overall, the right eye, the folds between both eyebrows, and the mouth areas were more informative with posed than with spontaneous and high-intensity spontaneous expressions. Note that the absence of a main effect of emotion does not indicate that the nature of the facial signal is the same across expressions; rather, it indicates that the signal is located in the same facial areas across expressions. For instance, the eye area may be used to recognize both disgust and surprise expressions, even though the eyes do not take on the same shape for both expressions.

Finally, in order to verify whether the increased intensity in high-intensity spontaneous expressions had an impact on the model observers' performance, we compared the signal-to-noise ratio needed to maintain the same performance as with spontaneous expressions.

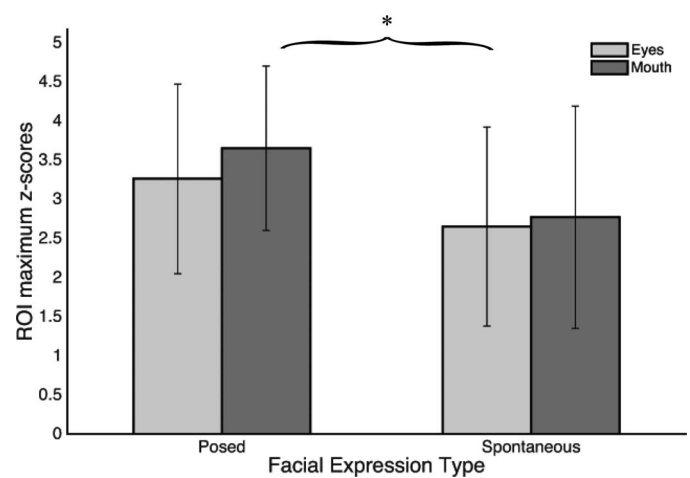


Figure 7. Average value of the maximum z scores reached in the mouth and eye areas for the model observers. Error bars represent standard deviation.

Since the same number of bubbles was used with both spontaneous and high-intensity spontaneous expressions, the amount of noise represented a direct measure of the task difficulty. The results indicate that a higher signal-to-noise ratio was needed with spontaneous expressions ($M = 0.94$, $SD = 0.06$) compared with high-intensity spontaneous ones ($M = 0.87$, $SD = 0.09$), $t(17) = 4.4$, $p < 0.001$, indicating that the model observers performed better with more intense spontaneous expressions.

Relative utilization of the mouth and eye areas

In order to gather a better understanding of the relative utilization of the mouth and eye areas, the same ROI analysis was performed as on the human data. A 2 (ROI: mouth vs. eyes) \times 2 (type of expression: posed vs. spontaneous) repeated-measures ANOVA was conducted on the maximum z -score values obtained in the ROI. The results are displayed in Figure 7. A main effect of the type of expression was obtained, $F(1, 19) = 25.6$, $p < 0.001$, indicating that on average the maximum z -score values were higher with posed expressions ($M = 3.46$, $SD = 1.13$) than with spontaneous ones ($M = 2.71$, $SD = 1.35$). The main effect of the ROI was marginally significant, $F(1, 19) = 4.4$, $p = 0.051$, indicating that on average the maximum z -score values were higher in the mouth area ($M = 3.21$, $SD = 1.33$) than the eye area ($M = 2.96$, $SD = 1.26$). However, in contrast with the human participants' results, the ROI \times Type of expression interaction was not significant, $F(1, 19) = 0.9$, $p = 0.368$, suggesting that the relative reliance on both areas did not differ for posed and spontaneous expressions.

Discussion

Research on the recognition of basic facial emotions has typically used images of posed facial expressions (e.g., Blais et al., 2012; Blais et al., 2017; Dailey et al., 2010; Ekman & Friesen, 1978; Elfenbein et al., 2007; Fiset et al., 2017; Smith et al., 2005; Sullivan et al., 2007; Thibault et al., 2012). However, the use of posed facial expressions rather than expressions resulting from genuine emotions has raised many questions about the ecological validity of these studies (Kayyal & Russell, 2013; Nelson & Russell, 2013; Russell, 1994), namely whether the findings could be generalized to the recognition of expressions occurring in our daily life. The present study is the first to directly compare the visual strategies underlying the recognition of spontaneous and posed facial expressions. Two key findings were obtained: Visual strategies are less systematic with spontaneous expressions (i.e., lower z scores than with posed expressions); and the relative utilization of the mouth and eyes to discriminate all four expressions from one another changed with spontaneous expressions.

Strategies used with spontaneous expressions are less systematic

Lower z scores were obtained in the CIs of spontaneous facial expressions than in those of posed ones, which most likely reflects a less systematic use of visual information with spontaneous than with posed expressions. This lower systematicity was also found for the model observers. The model observers provide a measure of the portion of results that may be explained by the properties of the stimuli and task and the portion that instead reflects constraints of the human visual system. The finding of lower z scores when the model observers were executed with spontaneous expressions therefore suggests that, overall, the signal was weaker in spontaneous than in posed expressions. However, our results indicate that not any kind of visual degradation leads to lower z scores. In fact, overall performance was maintained at the same level for spontaneous and posed expressions by applying a filter with a smaller number of bubbles on the latter, hence reducing the amount of facial information available. In a related vein, as for the model-observer analysis, more noise was applied to the high-intensity spontaneous stimuli than to the regular-intensity spontaneous ones, but the z scores did not significantly differ. Together, these results suggest that the lower z scores obtained with spontaneous expressions reflect the presence of a weaker signal in terms of the informativeness of the cues available.

The finding of a weaker signal in spontaneous than in posed expressions is not surprising. In fact, in day-

to-day social interactions, people sometimes modify their expression to hide the emotion they are truly experiencing. In their influential neuro-cultural theory of facial expressions, Ekman and Friesen (1971) proposed that depending on the context, an expression might be modified. It might be intensified, attenuated, neutralized, or even masked by being replaced by another expression. These modifications are more likely to occur with spontaneous than with posed expressions, since for posed expressions one is specifically asked to reproduce the expression that is normally associated with an emotion. In fact, the spontaneous expressions that were used in the present study were collected while individuals were viewing movies in an experimental setting; participants were thus aware of being filmed (Aifanti et al., 2010), which may have led them to slightly modify their expressions.

Indeed, as presented in the Method section, preliminary analyses conducted on the stimuli selected for the present study revealed that although the dominant emotion was perceived as more intense in the posed than in the spontaneous expressions, the emotion with the second highest rating was perceived as relatively more intense with spontaneous than with posed expressions. This suggests that the spontaneous expressions presented to participants were less intense, and more ambiguous, than the posed ones. These two factors may have contributed to the lower systematicity of the visual information used to perform the task, which was observed in both human and model observers.

Although the design of the present study did not allow us to quantify the relative contribution of intensity and ambiguity to the decreased systematicity of the visual strategies, we performed one additional analysis with the model observers to verify whether intensity by itself could explain the results. Specifically, the spontaneous stimuli selected for the experiment were used to generate new stimuli of this kind with a signal intensity that was increased linearly to 150% of that of the original ones. Then the model observers were run with this new set of stimuli. Indeed, this analysis demonstrated that the latter were more easily recognized by the model observers than the original spontaneous expressions. These results confirm the effectiveness of the intensity manipulation and suggest that the lower performance typically observed with spontaneous expressions may in part be explained by their lower intensity (Gan, Nie, Wang, & Ji, 2017; Kayyal & Russell, 2013). However, even with substantially more intense spontaneous stimuli, z scores were still significantly lower than with posed expressions, and did not significantly differ from those obtained with spontaneous expressions of lower intensity. This suggests that increasing the intensity of the spontaneous stimuli did not suffice to make the model observers' strategies more systematic, indicating

that the higher level of ambiguity observed with spontaneous expressions may significantly contribute to less systematic use of specific visual information.

Higher ambiguity may lead to a less systematic use of specific visual information in two different ways. First, it possibly entails more heterogeneity across participants. The analysis reported in the Results section indeed indicates that a slightly lower number of participants in the spontaneous condition than in the posed-expression condition adopted a strategy that overlapped with the average strategy. Another way in which high ambiguity may induce lower z scores in the CIs is by increasing the variability of information use across different exemplars of an expression. For instance, if in some pictures a disgusted expression was masked using a happy expression, one may succeed at interpreting the expression as reflecting disgust by taking into account the knowledge that people often mask disgust with happiness. If that is the case, then both the nasolabial folds and the corners of the mouth may lead to a correct answer. If, in addition, disgust were masked with an angry expression in other pictures, participants may again have used both the folds between the eyebrows and the nasolabial folds to correctly interpret the expression as disgust. This would result in a spreading of the potentially diagnostic information across the whole face, which would in turn reduce the z scores. The lower z scores obtained by the model observers with spontaneous compared to posed expressions suggest that the diagnostic information was indeed more spread out in the former than in the latter condition.

Spontaneous expressions being more ambiguous than posed ones is not a new finding. In fact, although very few studies have used spontaneous expressions to understand the processes underlying their recognition, the ones that have (Crivelli, Russell, Jarillo, & Fernández-Dols, 2017; Hess & Blair, 2001; Motley & Camden, 1988; Naab & Russell, 2007; Wagner, 1990; Wagner, MacDonald, & Manstead, 1986; Yik, Meng, & Russell, 1998; for a review, see Kayyal & Russell, 2013) have revealed a rather low level of endorsement of the predicted label for different exemplars of expressions. The present results support this finding, in that spontaneous expressions were perceived as containing more cues typically associated with other expressions. However, they also extend this finding by showing that spontaneous expressions are recognized using more heterogeneous visual strategies than posed expressions, potentially because of this higher level of ambiguity.

The relative utilization of the mouth and eyes changes with spontaneous expressions

In a previous study (Blais et al., 2012), we showed using a similar methodology as the one in the present

study that the diagnostic visual information to recognize facial expressions is not uniformly distributed across the face. In fact, when the task consists of categorizing the six basic facial expressions, the mouth is the most diagnostic area to discriminate all six expressions. Interestingly, in that study the participants' reliance on the mouth was even higher than that predicted by a model observer, which provided a more objective estimation of the distribution of the diagnostic information in facial expressions. This suggests that the high reliance on the mouth is not just a reflection of where the signal is in the facial expressions, but is in part linked to the constraints of the visual system and the visual representations used to recognize expressions. However, this result was obtained with posed facial expressions.

The present study shows that the reliance on the mouth area decreases when participants attempt to categorize spontaneous expressions. Indeed, although a higher reliance on the mouth than on the eye area was found with posed expressions, thus replicating the finding of Blais et al. (2012), there was no significant difference in the utilization of the mouth and eye areas with spontaneous expressions. Most interestingly, the model-observer analysis did not reveal such a shift in the relative utilization of eyes and mouth with spontaneous expressions. This suggests that the pattern of results obtained by the human observers does not simply reflect the relative informativeness of the eyes and mouth areas in both kinds of expressions: Although their relative informativeness was similar for posed and spontaneous expressions, participants relied proportionally less on the mouth with spontaneous expressions.

As mentioned in the Introduction, the area of the mouth is more susceptible of being voluntarily controlled than the area of the eyes (Ekman & Friesen, 2003). Thus, if one wants to get accurate information about the emotion felt by another, the eyes may be a more reliable source. On the other hand, from a signal-based point of view the mouth contains the most discriminant information. One possibility to explain the higher reliance on the mouth area with posed expressions and the reduced reliance on that area with spontaneous expressions is that the presence of signal ambiguity affects the relative weight allocated to the mouth and eye areas. With posed expressions, the signal ambiguity is very low; therefore, a high processing weight is attributed to the mouth area, since it is the most informative one to discriminate across all expressions. With spontaneous expressions, the signal ambiguity is higher; it may indicate an attempt to mask an expression, and thus the reliance on the mouth—which may in this case convey partly inaccurate information—decreases. Of course, this proposition would require empirical validation, for instance by

parametrically manipulating signal ambiguity while measuring the variations in the visual strategies used. Nevertheless, the present finding brings new nuances to the previous results obtained with posed expressions, and shows that while a high reliance on the mouth area is found with posed expressions, this does not appear to be the case with spontaneous ones.

Limits of the present study

In order for us to be able to compare visual strategies with posed and spontaneous expressions, a few constraints needed to be respected. First, we needed pictures of good quality in which the individuals were presented in a full frontal view. Moreover, to allow a more direct comparison of spontaneous and posed expressions, we needed pictures of the same individuals while they produced the two kinds of expressions. Such constraints made it almost impossible to have stimuli that would reflect spontaneous expressions captured outside of a laboratory context. Indeed, the stimuli used were recorded within a laboratory setting, while participants viewing the videos knew they were being filmed. The fact that the participants knew they were part of an experiment and that they were being filmed may have led them to control their facial expressions.

In addition, because of how the Bubbles method works, the selected expressions needed to be recognizable. On the one hand, this implies that the expressions selected were the ones that were the most obvious in the recordings of the MUG database. Studies that have previously assessed the recognizability of spontaneous expressions have revealed low levels of endorsement of predicted labels, ranging from 15% to 66% (for a review, see Kayyal & Russell, 2013). In the present study, the spontaneous expressions used led to a level of endorsement equivalent to the upper limit of the range observed previously (66.8%). On the other hand, as explained in the Method section, the Bubbles method required that the level of accuracy at recognizing the expressions needed to be quite high before bubbles could be applied on it. Otherwise, it would not have been possible to know whether a trial was failed because the unmasked emotion was unrecognizable or because the useful visual information was masked by the bubbles mask. Thus, this implies that the visual strategies revealed in the present study were measured with stimuli that were viewed several times during the experiment, and for which participants received training with respect to labeling. Although this may affect the ecological validity of the results, previous research does not support the idea that repetitive exposure to a limited number of stimuli induces biases in visual strategies. For instance, in one study the Bubbles method was used with famous faces, and each identity

was presented only once to each participant (Butler, Blais, Gosselin, Bub, & Fiset, 2010). The results of that study were highly consistent with previous studies in which each face identity was presented hundreds of times—namely, the eye region was revealed as the most diagnostic for face identification. Relatedly, another study (Royer et al., 2015) compared the performance of participants in a task with the Bubbles method with their performance in tasks with unaltered faces (e.g., Cambridge Face Memory Test). It was shown that the correlation increased as a function of the number of trials performed in the task with the Bubbles method, suggesting that the overlap between the mechanisms involved in the Bubbles task and the “normal” face-processing mechanisms actually increased as a function of exposure to the task. Similarly, Royer et al. (2016) showed that the reliance on the eye region increases as a function of the number of trials performed in a Bubbles task. Again, this suggests that the more the participants were exposed to the face identities (or the more familiar they became with the identities), the more their strategy became stable and focused on the information that was most diagnostic across a high number of identities. Together, these results suggest that the heavy training often found in Bubbles tasks does not induce biased strategies during face identification. Thus, it is safe to assume that the same would be true with facial expressions.

Another limit of the present study is that even though the selected facial expressions were evaluated by an independent sample of participants and reached a considerable level of endorsement, the MUG database did not provide any measure of the emotions actually felt by the individuals while they viewed the video. The spontaneous expressions used in the present study may therefore reflect different combination of emotions for different individuals. For instance, when they viewed disgusting scenes, even if disgust was the dominant expression, some individuals appeared partly amused, and others appeared partly shocked. This may have contributed to the higher heterogeneity observed in the visual strategies used with spontaneous expressions. It may be interesting, in a future study, to take into account felt emotions when measuring the visual strategies underlying spontaneous-expression recognition. Moreover, it may be interesting to verify the generalizability of the present results using a different facial-expression database. In fact, studies measuring the spatial facial variation in posed and spontaneous expressions using Bayesian networks (Wang, Wu, He, Wang, & Ji, 2015) and deep convolutional neural networks (Gan et al., 2017) have suggested that even though significant differences in the movement of some feature points are found between the two types of stimuli (i.e., mouth width, lip corner, and brows), those features tend to differ depending on the database that

was used. These results suggest that the spatial-information utilization might be sensitive to the database setup.

Despite the obvious limits that these constraints impose on the ecological validity of the present results, the spontaneous expressions we used were clearly induced in a more natural way than the posed facial expressions previously used in research about the visual processes underlying the recognition of facial expressions of emotions. Thus, the present study allows us to further our understanding of facial-expression recognition: It shows that with spontaneous expressions, the visual strategies are more heterogeneous and less processing weight is attributed to the mouth area.

Future studies should verify whether the heterogeneity observed in the visual strategies decreases with dynamic spontaneous expressions. In fact, previous studies have shown that the motion contained in dynamic expressions contains useful information that improves facial-expression categorization (Ambadar, Schooler, & Cohn, 2005; Cunningham & Wallraven, 2009a, 2009b; Chiller-Glaus, Schwaninger, Hofer, Kleiner, & Knappmeyer, 2011; Matsuzaki & Sato, 2008; but see Gold, 2013). Moreover, differences have been observed in the temporal unfolding of spontaneous and posed expressions (Ross et al., 2007; Ross & Pulusu, 2013). It is possible that motion is even more helpful when it comes to disambiguating spontaneous expressions. Motion may guide an observer with regard to which facial area is most likely to provide useful information to recognize an expression; if that is true, more homogeneity would be expected in the visual information used with spontaneous dynamic expressions. In a related vein, with dynamic posed expressions it has been shown that the mouth area was used earlier and for a longer duration than the eye area (Blais et al., 2012). Given that the mouth is more easily controlled to mask emotions, participants may show a change in the temporal unfolding of their information utilization with spontaneous expressions. For instance, they may decrease reliance on the mouth as a function of time to avoid using it once individuals have started controlling its appearance.

Conclusion

In past years, most research on the recognition of facial expressions has used posed rather than spontaneous facial expressions. The present study is the first to directly compare the visual strategies underlying the recognition of the two types of expressions. The results reveal that spontaneous expressions are recognized using more heterogeneous strategies. Moreover, the high reliance on the mouth observed

with posed expressions is not maintained with spontaneous expressions. This finding may indicate that when confronted with ambiguous facial cues, individuals decrease their utilization of the mouth area, which may contain inaccurate information regarding the emotion truly felt by someone. Although this was a first step toward a better understanding of facial-expression recognition in a more ecological setting, more research will be needed to fully understand the visual processes involved in the processing of dynamic, spontaneous expressions, as well as how the context in which the expressions occur influences these processes.

Keywords: facial expressions, perceptual strategies, psychophysics, facial features

Acknowledgments

We would like to thank the team of the MUG face database for kindly granting us access to their stimuli. This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada (CRSNG) to Caroline Blais, a graduate scholarship from the Fonds de Recherche du Québec pour la Nature et les Technologies (FRQNT) to Camille Saumure, and an undergraduate scholarship from CRSNG to Marie-Pier Plouffe-Demers.

Commercial relationships: none.

Corresponding author: Caroline Blais.

Email: caroline.blais@uqo.ca.

Address: Department of Psychoeducation and Psychology, Université du Québec en Outaouais, Gatineau, Québec, Canada.

References

- Adolphs, R. (2001). The neurobiology of social cognition. *Current Opinion in Neurobiology*, *11*(2), 231–239.
- 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.
- Aifanti, N., Papachristou, C., & Delopoulos, A. (2010). The MUG facial expression database. *Presented at the 2010 11th International Workshop on Image Analysis for Multimedia Interactive Services* (pp. 1–4). Desenzano del Garda, Brescia: IEEE. Retrieved from <https://ieeexplore.ieee.org/document/5617662/>

- 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.
- Blais, C., Fiset, D., Roy, C., Saumure, C., & Gosselin, F. (2017). Eye fixation patterns for categorizing static and dynamic facial expressions. *Emotion*, *17*(7), 1107–1119.
- Blais, C., Roy, C., Fiset, D., Arguin, M., Gosselin, F. (2012). The eyes are not the window to basic emotions. *Neuropsychologia*, *50*(12), 2830–2838.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.
- Burt, P., & Adelson, E. (1983). The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, *31*, 532–540.
- Butler, S., Blais, C., Gosselin, F., Bub, D., & Fiset, D. (2010). Recognizing famous people. *Attention, Perception, & Psychophysics*, *72*(6), 1444–1449.
- Chambon, V., & Baudouin, J. Y. (2009). Reconnaissance de l'émotion faciale et schizophrénie. *L'Évolution psychiatrique*, *74*(1), 123–135.
- Chauvin, A., Worsley, K. J., Schyns, P. G., Arguin, M., & Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *Journal of Vision*, *5*(9):1, 659–667, <https://doi.org/10.1167/5.9.1>. [PubMed] [Article]
- Chiller-Glaus, S. D., Schwaninger, A., Hofer, F., Kleiner, M., & Knappmeyer, B. (2011). Recognition of emotion in moving and static composite faces. *Swiss Journal of Psychology*, *70*, 233–240.
- Clark, C. M., Gosselin, F., & Goghari, V. M. (2013). Aberrant patterns of visual facial information usage in schizophrenia. *Journal of Abnormal Psychology*, *122*, 513–519.
- Crivelli, C., Russell, J. A., Jarillo, S., & Fernández-Dols, J. M. (2017). Recognizing spontaneous facial expressions of emotion in a small-scale society of Papua New Guinea. *Emotion*, *17*(2), 337–347.
- Cunningham, D. W., & Wallraven, C. (2009a). 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]
- Cunningham, D. W., & Wallraven, C. (2009b). The interaction between motion and form in expression recognition. In *Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization* (pp. 41–44). New York, NY: ACM.
- Dailey, M. N., Joyce, C., Lyons, M. J., Kamachi, M., Ishi, H., Gyoba, J., & Cottrell, G. W. (2010). Evidence and a computational explanation of cultural differences in facial expression recognition. *Emotion*, *10*(6), 874–893.
- Duncan, J., Gosselin, F., Cobarro, C., Dugas, G., Blais, C., & Fiset, D. (2017). Orientations for the successful categorization of facial expressions and their link with facial features. *Journal of Vision*, *17*(14):7, 1–16, <https://doi.org/10.1167/17.14.7>. [PubMed] [Article]
- Dunlap, K. (1927). The role of the eye muscles and mouth muscles in the expression of emotions. *Genetic Psychology Monographs*, *2*, 199–233.
- Eisenbarth, H., & Alpers, G. W. (2011). Happy mouth and sad eyes: Scanning emotional facial expressions. *Emotion*, *11*(4), 860–865.
- Ekman, P. (1972). Universals and cultural differences in facial expression of emotion. In J. Cole (Ed.), *Nebraska Symposium on Motivation* (pp. 207–283). Lincoln, NE: University of Nebraska Press.
- Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17*(2), 124–129.
- Ekman, P., & Friesen, W. (1978). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.
- Ekman, P., & Friesen, W. V. (2003). *Unmasking the face: A guide to recognizing emotions from facial clues*. Los Altos, CA: Malor Books.
- Elfenbein, H. A., Beaupré, M., Lévesque, M., & Hess, U. (2007). Toward a dialect theory: Cultural differences in the expression and recognition of posed facial expressions. *Emotion*, *7*(1), 131–146.
- 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.
- Gan, Q., Nie, S., Wang, S., & Ji, Q. (2017). Differentiating between posed and spontaneous expressions with latent regression Bayesian network. Presented at the *Thirty-First AAAI Conference on Artificial Intelligence* (pp. 4039–4045). San Francisco CA: AAAI Press. Retrieved from <https://www.aaai.org/ocs/index.php/AAAI/AAAI17>
- Gan, Q., Wu, C., Wang, S., & Ji, Q. (2015). Posed and spontaneous facial expression differentiation using deep Boltzmann machines. Presented at the *2015 International Conference on Affective Computing and Intelligent Interaction* (pp. 643–648). Xi'an, Shaanxi: IEEE.
- Gold, J. M., Barker, J. D., Barr, S., Bittner, J. L., Bromfield, W.D., Chu, W. D., ... & 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]
- Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research*, *41*(17), 2261–2271.
- Gosselin, F. & Schyns, P. G. (2002). RAP: A new framework for visual categorization. *Trends in Cognitive Science*, *6*, 70–77.
- Hall, J., Harris, J. M., Sprengelmeyer, R., Sprengelmeyer, A., Young, A. W., Santos, I. M., . . . Lawrie, S. M. (2004). Social cognition and face processing in schizophrenia. *The British Journal of Psychiatry*, *185*(2), 169–170.
- Harms, M. B., Martin, A., & Wallace, G. L. (2010). Facial emotion recognition in autism spectrum disorders: A review of behavioral and neuroimaging studies. *Neuropsychology Review*, *20*(3), 290–322.
- Hess, U., & Blairy, S. (2001). Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy. *International Journal of Psychophysiology*, *40*, 129–141.
- Hooker, C., & Park, S. (2002). Emotion processing and its relationship to social functioning in schizophrenia patients. *Psychiatry Research*, *112*(1), 41–50.
- Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. *Current Biology*, *19*(18), 1543–1548.
- Jack, R. E., & Schyns, P. G. (2017). Toward a social psychophysics of face communication. *Annual Review of Psychology*, *68*, 269–297.
- Kayyal, M. H., & Russell, J. A. (2013). Americans and Palestinians judge spontaneous facial expressions of emotion. *Emotion*, *13*(5), 891–904.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, *36*(14), 1–16.
- Kring, A. M., & Elis, O. (2013). Emotion deficits in people with schizophrenia. *Annual Review of Clinical Psychology*, *9*, 409–433.
- Krolak-Salmon, P. (2011). La reconnaissance des émotions dans les maladies neurodégénératives. *La Revue de médecine interne*, *32*(12), 721–723.
- Lee, J., Gosselin, F., Wynn, J. K., & Green, M. F. (2011). How do schizophrenia patients use visual information to decode facial emotion? *Schizophrenia Bulletin*, *37*(5), 1001–1008.
- Mandal, M. K., Pandey, R., & Prasad, A. B. (1998). Facial expressions of emotions and schizophrenia: A review. *Schizophrenia Bulletin*, *24*(3), 399–412.
- Matsumoto, D., Ollide, A., Schug, J., Willingham, B., & Callan, M. (2009). Cross-cultural judgments of spontaneous facial expressions of emotion. *Journal of Nonverbal Behavior*, *33*(4): 213.
- Matsuzaki, N., & Sato, T. (2008). The perception of facial expressions from two-frame apparent motion. *Perception*, *37*, 1560–1568.
- Motley, M., & Camden, C. (1988). Facial expression of emotion: A comparison of posed expressions versus spontaneous expressions in an interpersonal communication setting. *Western Journal of Speech Communication*, *52*, 1–22.
- Naab, P. J., & Russell, J. A. (2007). Judgments of emotion from spontaneous facial expressions of New Guineans. *Emotion*, *7*, 736–744.
- Nelson, N. L., & Russell, J. A. (2013). Universality revisited. *Emotion Review*, *5*(1), 8–15.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.
- Ross, E. D., Prodan, C. I., & Monnot, M. (2007). Human facial expressions are organized functionally across the upper-lower facial axis. *The Neuroscientist*, *13*(5), 433–446.
- Ross, E. D., & Pulusu, V. K. (2013). Posed versus spontaneous facial expressions are modulated by opposite cerebral hemispheres. *Cortex*, *49*(5), 1280–1291.
- Royer, J., Blais, C., Barnabé-Lortie, V., Carré, M., Leclerc, J., & Fiset, D. (2016). Efficient visual information for unfamiliar face matching despite viewpoint variations: It's not in the eyes! *Vision Research*, *123*, 33–40.
- Royer, J., Blais, C., Gosselin, F., Duncan, J., & Fiset, D. (2015). When less is more: Impact of face processing ability on recognition of visually degraded faces. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(5), 1179–1183.
- Russell, J. A. (1994). Is there universal recognition of emotion from facial expressions? A review of the cross-cultural studies. *Psychological Bulletin*, *115*(1), 102–141.
- Simoncelli, E. P. (1999). Image and multi-scale pyramid tools [Computer software]. New York, NY: Author.
- Smith, M., Cottrell, G., Gosselin, F., & Schyns, P. G. (2005). Transmitting and decoding facial expressions of emotions. *Psychological Science*, *16*, 184–189.

- Smith, M. L., & Merlusca, C. (2014). How task shapes the use of information during facial expression categorizations. *Emotion, 14*(3), 478–487.
- Spezio, M. L., Adolphs, R., Hurley, R. S., & Piven, J. (2007). Abnormal use of facial information in high-functioning autism. *Journal of Autism and Developmental Disorders, 37*(5), 929–939.
- Sullivan, S., Ruffman, T., & Hutton, S. B. (2007). Age differences in emotion recognition skills and the visual scanning of emotion faces. *Journal of Gerontology: Psychological Sciences, 62B*, 53–60.
- Thibault, P., Levesque, M., Gosselin, P., & Hess, U. (2012). The Duchenne marker is not a universal signal of smile authenticity—but it can be learned! *Social Psychology, 43*, 215–221.
- Wagner, H. L. (1990). The spontaneous facial expression of differential positive and negative emotions. *Motivation and Emotion, 14*, 27–43.
- Wagner, H., MacDonald, C., & Manstead, A. (1986). Communication of individual emotions by spontaneous facial expressions. *Journal of Personality and Social Psychology, 50*, 737–743.
- Wang, S., Wu, C., He, M., Wang, J., & Ji, Q. (2015). Posed and spontaneous expression recognition through modeling their spatial patterns. *Machine Vision and Applications, 26*(2–3), 219–231.
- Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics, 33*, 113–120.
- 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.
- Yik, M. S. M., Meng, Z., & Russell, J. A. (1998). Adults' freely produced emotion labels for babies' spontaneous facial expressions. *Cognition & Emotion, 12*, 723–730.