Emotion
2017, Vol. 17, No. 7, 1107–1119

1528-3542/17/$12.00 http://dx.doi.org/10.1037/emo0000283

Eye Fixation Patterns for Categorizing Static and Dynamic Facial Expressions

Caroline Blais and Daniel Fiset
Université du Québec en Outaouais

Cynthia Roy
Université de Montréal

Camille Saumure Régimbald
Université du Québec en Outaouais

Frédéric Gosselin
Université de Montréal

Facial expressions of emotion are dynamic in nature, but most studies on the visual strategies underlying the recognition of facial emotions have used static stimuli. The present study directly compared the visual strategies underlying the recognition of static and dynamic facial expressions using eye tracking and the Bubbles technique. The results revealed different eye fixation patterns with the 2 kinds of stimuli, with fewer fixations on the eye and mouth area during the recognition of dynamic than static expressions. However, these differences in eye fixations were not accompanied by any systematic differences in the facial information that was actually processed to recognize the expressions.

Keywords: facial expressions, dynamic, visual strategies, eye movements, bubbles

Supplemental materials: http://dx.doi.org/10.1037/emo0000283.supp

In day-to-day social interactions, humans must constantly monitor the facial expressions of others to consequently adapt their behavior. Facial expressions are dynamic in nature, and most of the time their appearance is brief before the face returns to a neutral expression. The visual information necessary to recognize the emotion expressed in a face must therefore be extracted quickly.

Until recently, most studies have used static facial expressions in their quest of the strategies used by the visual system to extract the information necessary for the recognition of each basic emotion, mainly because technologies were more limited at the time this issue started to be studied. Many of these studies focused on either the ocular fixation patterns underlying the extraction of the visual information necessary for expression recognition or the visual information that is actually extracted and processed and that leads to a successful recognition. When ocular fixation patterns are measured, the results typically indicate a higher proportion of fixations on the mouth and the eye areas (Beaudry, Roy-Charland, Perron, Cormier, & Tapp, 2014; Eisenbarth & Alpers, 2011; Jack, Blais, Scheepers, Schyns, & Caldara, 2009; Schurgin et al., 2014; Vaidya, Jin, & Fellows, 2014). If the task is a classic categorization of basic facial expressions at their apex, the relative proportion of fixation on the different face areas is not modulated by the expression processed (Jack et al., 2009; Vaidya et al., 2014), although the mouth area is sometimes more fixated during the recognition of happiness and the eye area is sometimes more fixated during the recognition of sadness (Beaudry et al., 2014; Eisenbarth & Alpers, 2011). However, the pattern becomes more differentiated across facial expressions with subtle facial expressions (Vaidya et al., 2014) or when the task requires searching for the presence of a specific expression (Schurgin et al., 2014). On the other hand, when the visual information utilization is assessed, the results indicate different patterns for the six basic emotions: The mouth is the most useful area for the recognition of happiness and surprise, the eyes area is utilized for the recognition of anger and fear, the nose wrinkles and nasolabial folds are used for the recognition of disgust, and the eyebrows and mouth are used for the recognition of sadness (M. L. Smith, Cottrell, Gosselin, & Schyns, 2005).

However, recent research has suggested that dynamic and static facial expressions may be processed differently. In fact, it has been

This article was published Online First April 3, 2017.

Caroline Blais and Daniel Fiset, Département de psychoéducation et de psychologie, Université du Québec en Outaouais; Cynthia Roy, Centre de Recherche en Neuropsychologie et Cognition, Département de Psychologie, Université de Montréal; Camille Saumure Régimbald, Département de psychoéducation et de psychologie, Université du Québec en Outaouais; Frédéric Gosselin, Centre de Recherche en Neuropsychologie et Cognition, Département de Psychologie, Université de Montréal.

This work was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Grants RGPIN/04777–2014 to Frédéric Gosselin, 2108640 to Caroline Blais, and 1607508 to Daniel Fiset and by a graduate scholarship from the Fonds de Recherche Québécois en Nature et Technologies to Caroline Blais and a graduate scholarship from NSERC to Cynthia Roy. We thank Lao Junpeng, from the University of Fribourg, for his helpful suggestions with the eye-tracking analyses and the use of *i*Map4.

Correspondence concerning this article should be addressed to either Caroline Blais, Département de psychoéducation et de psychologie, Université du Québec en Outaouais, C.P. 1250, Succ. Hull, Gatineau, Québec, Canada J8X 3X7, or to Frédéric Gosselin, Département de psychologie, Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal, Québec, Canada H3C 3J7. E-mail: caroline.blais@uqo.ca or frederic.gosselin@umontreal.ca

shown that they are processed in partially different brain structures (Humphreys, Donnelly, & Riddoch, 1993; Schultz & Pilz, 2009) and that dynamic expressions lead to a greater activation of many structures known for their involvement in facial expression processing (Kilts, Egan, Gideon, Ely, & Hoffman, 2003; LaBar, Crupain, Voyvodic, & McCarthy, 2003; Sato, Kochiyama, Yoshikawa, Naito, & Matsumura, 2004; Trautmann, Fehr, & Herrmann, 2009). Moreover, many studies have revealed a different performance with dynamic compared to static facial expressions (Ambadar, Schooler, & Cohn, 2005; Chiller-Glaus, Schwaninger, Hofer, Kleiner, & Knappmeyer, 2011; Cunningham & Wallraven, 2009; Hammal, Gosselin, & Fortin, 2009; Matsuzaki & Sato, 2008). Thus, dynamic facial expressions have gradually started to be used in studies aiming to reveal the visual strategies underlying facial emotion recognition.

A few studies have measured the ocular fixation patterns occurring during the processing of visual scenes containing dynamic facial stimuli (Coutrot & Guyader, 2014; Võ, Smith, Mital, & Henderson, 2012) or during the processing of dynamic facial expressions (Buchan, Paré, & Munhall, 2007; Lischke et al., 2012). However, the stimuli used in these studies had a long duration and did not present the ocular fixation pattern specifically occurring while the emotion was expressed. Moreover, the eye movements occurring during the recognition of static and dynamic facial expressions have, to the best of our knowledge, never been directly compared using the same set of actors.

There have also been a few studies that measured the facial information and/or motion utilization during the processing of dynamic facial expressions. Blais, Roy, Fiset, Arguin, and Gosselin (2012) showed, using the Bubbles technique, that the mouth is the most crucial area to discriminate across all basic emotions, for both static and dynamic expressions. We completely reanalyze the data from this study in Experiment 2 of the current study. Nusseck, Cunningham, Wallraven, and Bülthoff (2008) investigated which facial area contains motion that is necessary and/or sufficient for the recognition of basic facial expressions as well as complex ones. They did so by specifically freezing the motion of some facial areas while the rest of the face moved naturally, and they observed the impact of such a manipulation on recognition. Their results showed that the facial areas or combinations of facial areas that led to recognition varied across expressions and were congruent, although not perfectly overlapping, with those found in previous studies using static expressions (see, e.g., M. L. Smith et al., 2005). More recently, Yu, Garrod, and Schyns (2012; see also Yu, Garrod, Jack, & Schyns, 2015) created a computational method to artificially synthesize facial expressions in which the different action units of the Facial Action Coding System (FACS; Ekman & Friesen, 1978) can be dynamically activated with varying temporal profiles. This method is akin to the reverse correlation method (Ahumada & Lovell, 1971), in that the different action units and their temporal profile are randomly selected on each trial and the participant's task is to decide which emotion (Jack, Garrod, & Schyns, 2014; Jack, Garrod, Yu, Caldara, & Schyns, 2012; Jack, Sun, Delis, Garrod, & Schyns, 2016; Yu et al., 2012), among a set of predetermined answers, the expressions are most similar to. Thus, it allows revealing the mental representation that someone has of a facial expression (i.e., which action units are coded in the visual memory that someone has of an expression). Using that method, they revealed which action units are correlated with the perception of the six basic facial expressions, as well as their temporal profile. The action units that were found as being part of the representation of each facial expression were relatively similar to those proposed by Ekman and Friesen (1978) in the FACS (i.e., rho varying between .32 and .91 across expressions). There were, however, some notable differences between the action units proposed by Ekman and Friesen as being part of each expression and the one revealed with Yu et al.'s (2012) method (e.g., for anger, the brow lowerer was less activated and the nose wrinkle was more activated in Yu et al.'s, 2012, compared to Ekman & Friesen's, 1978). These differences are likely to come from the different kinds of information provided by Ekman and Friesen's work and by the reverse correlation method. In fact, whereas Ekman and Friesen focused on the information available in each facial expression, Yu et al. focused on the information that was actually coded in the visual memory of the participants.

Thus, although the aforementioned studies have provided inestimable information about the visual strategies for recognizing dynamic facial expressions, there is a gap in the literature if one wants to understand how static and dynamic facial expressions differ with regard to these strategies. In fact, to the best of our knowledge, the visual strategies used to recognize static and dynamic facial expressions have never been directly compared using the same set of actor stimuli, either in studies assessing the diagnostic information or in studies measuring fixation patterns. Any divergence observed between the visual strategies used to recognize static and dynamic facial expressions may come from some differences in the way the actors express the different emotions.

The three experiments presented in this article aimed to directly compare the visual strategies used during the recognition of static and dynamic facial emotions expressed by the same set of actors. In the first and main experiment, the ocular fixations patterns occurring during the recognition of static and dynamic facial expressions are measured. We discovered that observers spend less time fixating on the eye and mouth areas during the recognition of dynamic than during the recognition of static facial expression of emotions. In Experiment 2 we eliminate the diagnostic information for recognizing static and dynamic facial expressions using the Bubbles method (Gosselin & Schyns, 2001) as a potential explanation. Finally, in Experiment 3, we show that the visual strategy revealed in Experiment 2 with dynamic stimuli was not altered by the presence of bubbles on the stimuli.

Experiment 1

Method

Experiment 1 was divided into two tasks that measured the ocular fixation pattern occurring during the categorization of, respectively, static and dynamic emotional expressions.

Participants. Twenty Caucasian participants (10 male; mean age = 24.8 years) took part in both tasks. The sample size was selected based on previous studies using eye tracking and revealing differences in the patterns of eye fixations using around 20 participants (e.g., Blais, Jack, Scheepers, Fiset, & Caldara, 2008; Jack et al., 2009). All participants had normal or corrected-to-normal visual acuity. All procedures were carried out with the ethics approval of the Université de Montréal.

Materials and stimuli. Stimuli were displayed on a calibrated high-resolution cathode ray tube (CRT) monitor with a refresh rate of 60 Hz. The experimental program was written in Matlab (MATLAB and Statistics Toolbox Release, 2006), using functions from the Psychophysics Toolbox (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) and EyeLink Toolbox (Cornelissen, Peters, & Palmer, 2002). Eye movements and fixations were measured and recorded at 500 Hz with the head-mounted oculomotor system EyeLink II (SR Research, Hamilton, Ontario, Canada). Only the dominant eye was tracked, but viewing was binocular.

The stimuli consisted in 10 actors (five male) expressing the six basic emotions (i.e., anger, disgust, fear, happiness, sadness, and surprise) as well as the expressions of pain or neutrality. The analyses on the pain expression are not presented here (see however C. Roy et al., 2015), but note that the pattern of results presented later was the same whether the trials containing the pain expression were included or not. The pictures and videos were taken from the STOIC database (S. Roy et al., 2007). The dynamic version of the stimuli consisted in 500-ms videos starting with an expression of neutrality and ending at the apex of the expression. The main facial features (eyes, nose, mouth) were aligned across facial expressions and across actors using linear manipulations such as translation, rotation, and scaling. They were also aligned across temporal frames to minimize head movements in the stimuli. The static version of the stimuli consisted in the last frame of the videos, that is, the apex of the expression. All stimuli were gray-scaled, and their luminance was normalized. The face width subtended 5.72 degrees of visual angle at a viewing distance of 109 cm.

Procedure. Each participant completed the static and dynamic tasks in a counterbalanced order. Each task included two blocks of 80 trials, in which each of the 80 stimuli (i.e., 10 actors, eight expressions) was displayed once. The stimuli presented on each trial varied in a random order across participants. A nine-point calibration was performed with the eye tracker at the beginning of each block and repeated after 50 trials. A drift correction was also performed every five trials.

Each trial started with the presentation of a fixation cross, displayed in the middle of a uniform midgray screen for 500 ms. The fixation cross was then immediately replaced by a facial expression stimulus, displayed in the middle of a uniform midgray screen for 500 ms. Finally, the stimulus was replaced by a uniform midgray screen until the participant responded. The participants' task was to categorize aloud, among the eight possible choices, the facial expression presented. The experimenter pressed the key corresponding to the emotion perceived. No performance feedback was provided.

Results and Discussion

On average, participants correctly categorized the static expressions on 74.2% (SD = 24.0) of the trials and the dynamic expressions on 82.8% (SD = 15.2) of the trials. The average accuracy rate with each facial expression in the static and dynamic tasks is presented in Table 1. A repeated-measures analysis of variance (ANOVA) was conducted on the task and facial expressions variables, on the arcsine-transformed accuracy rates (e.g., Field, 2013). The result showed significant main effects of task, $F(1, 19) = 5.54$, $p = .029$, $\eta_p^2 = .23$, and facial expressions, $F(6, 114) = 28.0$, $p < .001$, $\eta_p^2 = .596$. The interaction did not reach significance, $F(6, 114) = 1.2$, $p = .309$, $\eta_p^2 = .06$.

Table 1
Average Accuracy Rates for the Static and Dynamic Facial Expressions in Experiment 1

Expression	Static	Dynamic
Anger	83.8 (22.5)	91.8 (10.0)
Disgust	70.8 (21.8)	83.5 (11.3)
Fear	58.5 (24.1)	75.8 (16.6)
Happiness	90.3 (22.6)	97.0 (7.1)
Neutrality	66.8 (18.8)	73.8 (7.8)
Pain	67.8 (26.4)	72.5 (21.6)
Sadness	73 (20.6)	79.5 (12.1)
Surprise	82.8 (21.3)	88.8 (10.6)

Note. Data are given as means, with standard deviations in parentheses.

Fixations and saccades were determined using a custom algorithm, that is, a saccade velocity threshold of 30°/s and a saccade acceleration threshold of 4000°/s^2. The following statistics include the first fixation. On average, participants made 2.11 fixations per trial (SD = .27) in the static condition, with an average fixation duration of 223.7 ms (SD = 45.4). In the dynamic condition, they made on average 2.01 fixations per trial (SD = .29), with an average fixation duration of 251.8 ms (SD = 60.5). There was no significant difference between the two conditions on the number of fixations, $t(19) = 1.36$, $p = .19$, but a marginal trend for longer fixations was observed with dynamic than with static expressions, $t(19) = 1.88$, $p = .08$.

Maps representing the proportion of time spent fixating across the face were computed for each emotion in each task. Only the fixations, including the first one, were entered in the computation of these maps. More specifically, the maps were computed by calculating, for each pixel of a stimulus image, the total amount of time it was fixated across all trials, divided by the total amount of time all the pixels in the image were fixated across trials. These maps therefore represented pixel-based proportions of fixation time on the stimuli. The maps were smoothed using a Gaussian kernel with a full width at half maximum of 33 pixels (1° of visual angle). Statistical differences across the emotions and across the tasks were tested using the *i*Map toolbox, Version 4 (Lao, Miellet, Pernet, Sokhn, & Caldara, 2016). This toolbox applies univariate pixelwise linear mixed models on the fixation maps instead of dividing it into regions of interest and uses a bootstrap procedure to correct for multiple comparisons. Thus, using iMap4, a 7 (facial expressions) * 2 (static or dynamic condition) repeated-measures ANOVA was performed across all pixels using Linear Mixed Model. The pixelwise ANOVAs result in a map of F values for each effect (e.g., effect of the emotion, effect of the condition, Emotion * Condition interaction), and the bootstrap procedure aims at finding the significant clusters by comparing the cluster characteristic (i.e., cluster mass) with the null distribution computed from the centered data. (see Lao et al., 2016, for technical details on the procedure).

The main effect of emotion was not significant, indicating that the fixation pattern did not differ across facial expressions. There was, however, a significant effect of the task, showing that the pattern of fixations differed with the static and the dynamic facial expressions. In fact, fixations were focused on the center of the face more in the dynamic than in the static condition and focused on the left eye and mouth more in the static than in the dynamic condition. The interaction between the emotion presented and the task was not significant. Note that on average, eyes and mouth were, respectively, 1.7° and 2.1° of visual angle apart from the center of the face (initial fixation location). Figure 1 shows the areas that significantly differed between the fixation maps obtained with static and dynamic facial expressions. Figure 2 shows how stable this pattern was across expressions.

In summary, Experiment 1 revealed a different pattern of eye fixations for static and dynamic facial expressions, whereby the left eye and mouth are fixated more with static than with dynamic expressions and the middle of the face is fixated more with dynamic than with static faces. Such a difference in the pattern of fixations may reflect the utilization of different facial information, for instance a reliance on visual information contained in the eyes and mouth areas that is higher with static than with dynamic expressions. However, because it is possible to process information peripheral to the fixation location, it is also possible that the same facial areas are processed with dynamic and static faces, despite different eye fixation patterns. The aim of Experiment 2 was to measure the visual information utilization during the categorization of static and dynamic facial expressions. This would allow for drawing a more complete picture of the strategies underlying the processing of static and dynamic expressions, one that includes information regarding both what facial areas are fixated and what facial areas are actually processed and lead to a successful categorization.

Experiment 2

Method

Experiment 2 was divided in two tasks: The first task addressed information utilization during the recognition of static facial expressions, and the second task addressed information utilization during the recognition of dynamic facial expressions. The Bubbles method (Gosselin & Schyns, 2001) was used to address these questions. The data collected in these two experiments were part of a larger project in which we explored many dimensions of information utilization for the recognition of basic facial expressions. A series of analyses on these data have already been published (Blais et al., 2012). The aim of this former study was to verify whether, as often implied in popular culture as well as in the scientific literature, the eyes are the most crucial area for the recognition of emotions. The results showed that although the eye area is useful for emotion recognition, the mouth is the most diagnostic in discriminating all basic facial expressions from one another, whether they are static or dynamic. Thus, Blais et al. (2012) did not systematically compare the information used for categorizing static and dynamic expressions, and they collapsed all the facial expressions and spatial frequency bands together. The present work provides a series of analyses specifically designed to compare each static and dynamic facial expression on each spatial frequency band and on each facial area (see Blais et al., 2012, for more technical details on the method).

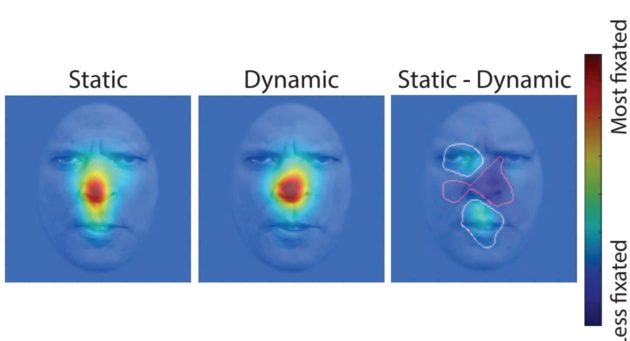

Figure 1. Maps of fixation durations in the static and dynamic conditions, as well as the difference between these maps. The areas that were fixated significantly more in the static condition are outlined in white, and the area that was fixated significantly more in the dynamic condition is outlined in red.

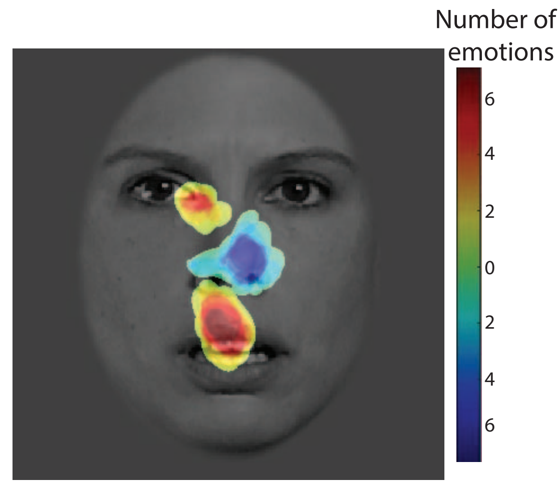

Figure 2. Stability across emotions of the difference between the maps of fixation durations in the static and in the dynamic conditions. Yellow–red areas show the number of emotions for which a higher density of fixations was obtained with static than with dynamic expressions. Green–blue areas show the number of emotions for which a higher density of fixations was obtained with dynamic than with static expressions.

Participants. Forty-one Caucasian participants (14 male; mean age = 24.2 years) took part in the task with static facial expressions, and 59 different Caucasian participants (30 male; mean age = 23.9 years) took part in the task with dynamic facial expressions. The number of participants and trials was chosen based on the study by M. L. Smith et al. (2005), which included only a static condition. Because Blais et al. (2012) decided to adjust the number of bubbles on the average accuracy across all emotions (to prevent the number of bubbles from becoming a cue for the emotion) rather than on the individual accuracy for each emotion as did M. L. Smith et al. and because they decided to test more participants who each performed fewer trials (to increase representativeness mostly), they increased the total number of trials by about 20%. There was no directly comparable precedent for the dynamic emotion condition, so the sample size criterion was estimated from a pair of studies on the identification of faces: one static, by Gosselin and Schyns (2001), and the other dynamic, by Vinette, Gosselin, and Schyns (2004). All participants had normal or corrected-to-normal visual acuity. All procedures were carried out with the ethics approval of the Université de Montréal.

Materials and stimuli. Stimuli were displayed on a calibrated high-resolution CRT monitor with a refresh rate of 60 Hz. The experimental program was written in Matlab, using functions from the Psychophysics Toolbox. The facial expression stimuli were the same as in Experiment 1. The face width subtended 5.72° of visual angle at an approximate viewing distance of 73 cm.

In the task with static facial expressions, a static version of Bubbles was used (Gosselin & Schyns, 2001). This technique consists in randomly sampling visual information on the space (i.e., *x, y* coordinates) and spatial frequency dimensions. Thus, on each trial, different subsets of facial areas filtered in five spatial frequency bands are presented to the participants. The general idea underlying the method is that if the facial areas needed for a successful expression categorization are available and encoded in the critical spatial frequencies, then the probability that the participant will answer correctly should increase. In contrast, if the diagnostic facial areas are not available or are not encoded in the right spatial frequencies, then the probability that the participant will answer correctly should decrease. In other words, by varying the available information across trials, it is possible to infer what facial information is used by participants to resolve the task. To use accuracy as a predictor variable in a regression, one must avoid floor and ceiling effects. Here, this was achieved by manipulating the number of bubbles to maintain the average accuracy rate halfway between chance and ceiling levels. Figure 3 provides a few examples of stimuli sampled using this version of the Bubbles method.

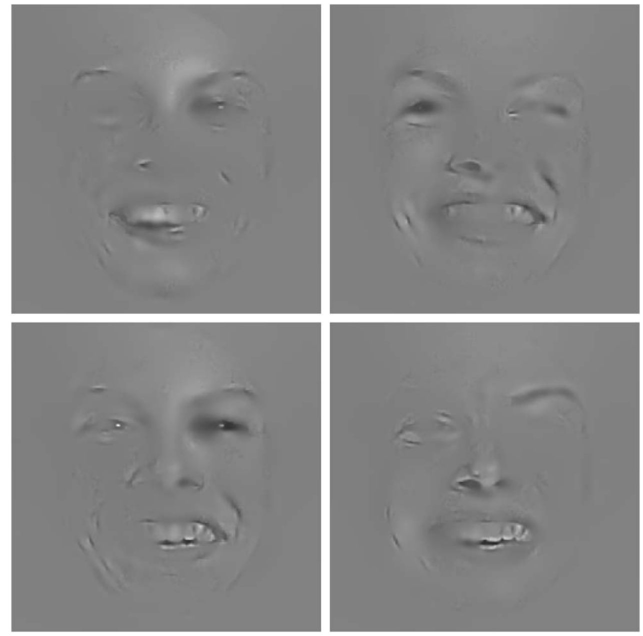

Figure 3. Four examples of stimuli created using the static Bubbles technique. The number of bubbles used to create these examples was the average number of bubbles needed by the participants of Experiment 2 to maintain the target accuracy rate. Note that the facial areas revealed, and the spatial frequency bands in which a facial area is revealed, vary from one trial to the other. For instance, the right eye is not revealed on the stimulus displayed in the bottom right panel, is revealed in the lower spatial frequency bands on the stimulus displayed in the bottom left panel, and is revealed in low as well as in high spatial frequencies on the stimulus displayed in the upper left panel.

In the task with the dynamic stimuli, a dynamic version of Bubbles was used (Blais et al., 2012; Vinette et al., 2004). Visual information was randomly sampled on the space (i.e., *x, y* coordinates), spatial frequency, and time dimensions. The creation of a stimulus on each trial is similar to the one with static bubbles except that, in addition to location and spatial frequency, time is sampled. The duration of each random sample depends on its spatial frequency. More specifically, the full width at half maximum of one bubble lasts 7.3, 6.1, 5.1, 4.2, and 3.5 frames from the highest to the lowest spatial frequency band. This was done to account for the faster processing of lower spatial frequencies (e.g., Hughes, Fendrich, & Reuter-Lorenz, 1990; Parker, Lishman, & Hughes, 1992). Thus, the general idea of the method is that if the facial areas needed for a successful expression categorization are encoded in the critical spatial frequencies and presented at the right moment during the stimulus processing, then the probability that the participant will answer correctly should increase. In contrast, if the diagnostic facial areas are not available, are not encoded in the right spatial frequencies, or are not presented at the right moment during the stimulus processing, then the probability that the participant will answer correctly should decrease. It is therefore possible to infer the temporal deployment of visual information utilization during the processing of dynamic facial expressions. In the analyses presented later, we focus on the comparison between the visual information used with static and dynamic faces and the temporal dimension is not explored, because it was not measured with static expressions (see Blais et al., 2012, for an analysis of the utilization of the facial features across time with dynamic expressions). Note that although the sampling space only partly overlapped for static and dynamic stimuli, previous studies have suggested that this has little impact on the results computed in the overlapping sampling subspaces. For instance, Dupuis-Roy, Dufresne, Fiset, and Gosselin (2012) sampled location, time, chrominance, and luminance during a face gender discrimination and obtained results in the location 2D subspace that replicated remarkably well those they had obtained previously when they had sampled only location (Dupuis-Roy, Fortin, Fiset, & Gosselin, 2009). Moreover, Vinette et al. (2004) sampled location and time during a face identification task and obtained results similar to the ones obtained by Gosselin and Schyns (2001), who sampled only location and spatial frequencies.

Procedure. The procedure was the same for both tasks. Each participant completed 4,000 trials, divided into experimental blocks of 160 trials. Most participants needed two experimental sessions (i.e., two separate days, not necessarily consecutive) to complete the task. Each trial started with the presentation of a fixation cross, displayed in the middle of a uniform midgray screen for 200 ms. The fixation cross was then immediately replaced by the stimulus, displayed in the middle of a uniform midgray screen for 500 ms. Finally, the stimulus was replaced by a uniform midgray screen until the participant responded. The participants' task was to categorize, among the eight possible choices, the facial expression presented. They indicated their response by pressing on the corresponding key on the keyboard. No performance feedback was provided. Accuracy rate was maintained at 56% correct (i.e., halfway between chance and perfect accuracy) on average across all expressions by adjusting the total number of bubbles on the stimulus on a trial-by-trial basis using QUEST (Watson & Pelli, 1983).

Results and Discussion

The average number of bubbles needed by the participants to maintain their mean accuracy rate around 56% by QUEST was 144.3 (SD = 119.6) and 241.3 (SD = 253.6) with static and dynamic stimuli, respectively. Their overall average accuracy rate was 61.8% for static expressions and 62.0% for dynamic faces. Their average accuracy rates per facial expression, in the static and dynamic tasks, are presented in Table 2.

To reveal the diagnostic information for recognizing static and dynamic basic facial expressions, we produced classification images using a procedure that amounts to a multiple linear regression on the bubbles masks (explanatory variable) and the participant's accuracy (predictor variable). This procedure, described in detail in Blais et al. (2012), results in z scores representing how strongly the processing of different areas of a face are related to the participants' accuracy. Collectively these z scores are called a classification image. Thus, the higher the absolute value of the z score is, the stronger the link between the processing of the facial area and the participants' accuracy is, and z scores close to zero indicate the absence of a link between the processing of the facial area and the participants' accuracy. The statistical significance of smooth classification images is often assessed using the Pixel Test or the Cluster Test from the Stat4Ci toolbox (Chauvin, Worsley, Schyns, Arguin, & Gosselin, 2005). Both of these tests correct for multiple comparisons by controlling the family-wise error rate, while taking into account the fact that contiguous pixels are not independent. The Pixel Test computes a statistical threshold based on the probability of observing a single pixel above the threshold. This test has been shown to be best suited for detecting focal signals with high z scores. The Cluster Test is more sensitive for detecting wide regions of contiguous pixels with relatively low z scores. It is based on the probability that, above a relatively low threshold t (e.g., 2.5), a cluster of size K (or more) pixels has occurred by chance. The two tests sometimes identify different statistically significant regions in smooth classification images.

Figure 4 shows the classification images obtained with static and dynamic facial expressions when averaged across all expressions and across all spatial frequencies. The areas depicted in color

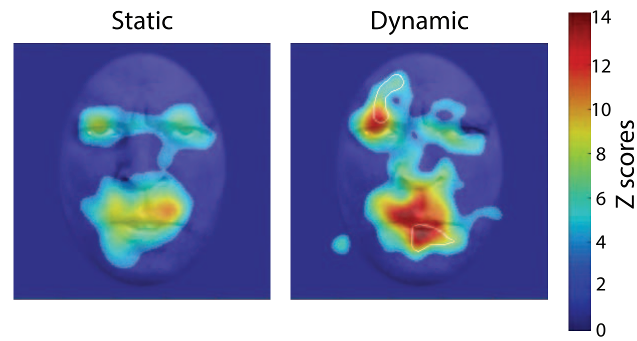

Figure 4. Maps of visual information significantly correlated with accuracy in recognizing static and dynamic facial expressions. The areas outlined in white correspond to the information that was significantly more useful in one condition than in the other.

were significantly linked with accuracy in categorizing static and dynamic facial expressions (Cluster Test; t = 3.5, K = 26.4, p < .05; maximum cluster size observed: 409.4 pixels). The areas outlined in white on the dynamic classification image were significantly more useful to categorize dynamic than static facial expressions (cluster test; t = 3.5, K = 123.8, p < .05; maximum cluster size observed: 506.8 pixels). No area was significantly more useful for categorizing static than dynamic facial expressions.

To quantify the overlap between the information used for static and that used for dynamic facial expressions, we calculated the proportion of significant pixels that overlapped in the static and dynamic conditions (i.e., number of significant pixels overlapping in the static and dynamic conditions, divided by the average of the number of significant pixels across both conditions). Note that for the purpose of this analysis, the pixel test was used to find the significant pixels. This is because the cluster test does not allow for concluding that every pixel within a significant cluster of pixels is significant. Among the pixels significantly useful with either dynamic or static stimuli, 74% overlapped across the two kinds of stimuli.

Even if the classification image representing the information used to categorize across all static expressions is similar to the one representing the information used to categorize across all the dynamic ones, it remains possible that some differences arise when each emotional expression is considered separately. A second analysis was therefore performed to compare the facial areas used to categorize each static and dynamic expression. Cluster tests were applied to find the clusters of pixels that reached a significance threshold in the static and dynamic classification images (t = 3.5, Bonferroni corrected across expressions; K = 297.7, p < .05; maximum cluster size observed: 680.7 pixels), as well as the clusters of pixels that were significantly more useful in one condition than in the other (t = 3.5, p < .05, Bonferroni corrected across expressions; K = 393.5; maximum cluster size observed: 776.5 pixels). Figure 5 shows the facial areas that were significantly linked with the accuracy in categorizing each static and dynamic expression. Among these areas, the ones that were significantly more useful in one condition than in the other are outlined in white or red. A significant difference could be found because of quantitative differences between the strategies used

Table 2
Average Accuracy Rates for the Static and Dynamic Facial Expressions in Experiment 2

Expression	Static	Dynamic
Anger	62.0 (11.5)	58.4 (9.2)
Disgust	45.7 (12.8)	47.4 (12.1)
Fear	57.7 (12.7)	57.1 (12.6)
Happiness	86.6 (8.1)	84.2 (7.4)
Neutrality	58.0 (11.3)	59.1 (9.2)
Pain	49.9 (13.9)	57.1 (14.4)
Sadness	63.8 (12.8)	56.0 (9.6)
Surprise	70.4 (12.2)	76.3 (11.3)

Note. Data are given as means, with standard deviations in parentheses.

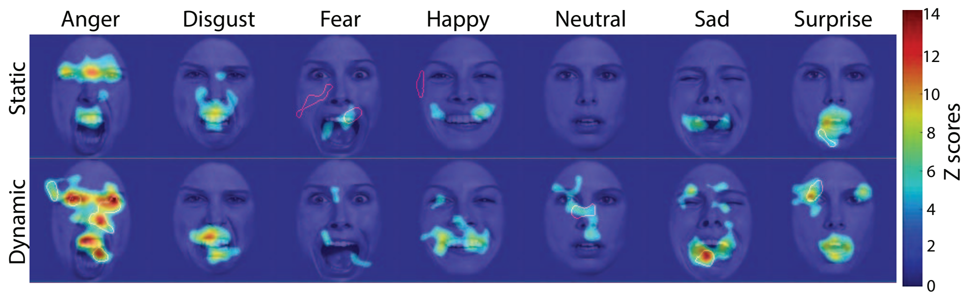

Figure 5. Maps of visual information significantly correlated with accuracy in recognizing each expression in its static and dynamic versions. The areas outlined in white correspond to the information that was significantly more useful in one condition than in the other because of quantitative differences, and those outlined in red correspond to the information that was significantly more useful in one condition than in the other because of qualitative differences between the strategies.

with static and dynamic stimuli; in that case, the area should be positively correlated with accuracy for both kinds of stimuli, although more correlated with one than the other. The areas corresponding to this scenario are depicted in white in Figure 5. On the other hand, a significant difference could also be found because of qualitative differences between the strategies used with static and dynamic stimuli; in that case, the area should be positively correlated with accuracy for one kind of stimulus but not correlated or negatively correlated with the other kind of stimulus (i.e., we used a criterion in which the z scores had to be equal to or below 0 with one kind of stimulus and positive with the other). The areas corresponding to this case scenario are depicted in red in Figure 5. Note that areas outlined on the classification images of static stimuli were more useful for static than for dynamic stimuli, and the areas outlined on the classification images of dynamic stimuli were more useful for dynamic than for static stimuli.

The overlap between static and dynamic information utilization was also calculated for each expression separately. Among the pixels significantly useful with either a dynamic or a static expression, an overlap of 72.5%, 46.3%, 58.0%, 48.6%, 44.6%, and 43.1% was found for anger, disgust, fear, happiness, sadness, and surprise, respectively. For the anger expression, the eyes area, the eyebrows, and the mouth were used with both static and dynamic expressions. With the disgust, fear, happiness, sadness, and surprise expressions, the mouth was used in both conditions. There were a few statistically significant differences between the static and dynamic expressions at the quantitative level: (a) The eyes area was more useful in the dynamic than in the static condition for the anger and surprise expressions, (b) the mouth area was more useful in the dynamic than in the static condition for the anger and sadness expressions, and (c) the mouth area was more useful in the static than in the dynamic condition for the surprise expression. There were also some statistically significant differences at the qualitative level: (a) A part of the cheek and the upper area of the mouth were useful in the static but not in the dynamic condition for the fear expression; (b) a part of the face contour close to the left eye area was useful in the happy static condition but not in the happy dynamic condition. Note that although the cheek and face contour were significantly more useful to categorize static than dynamic fear and happiness expressions, they were actually not useful for the recognition of static fear and happiness expressions (i.e., they did not reach the significance threshold). The significant difference in that specific case indicates that the cheek and face contour were positively, although moderately, correlated with accuracy, whereas these same areas were negatively, but again moderately, correlated with accuracy. Because these regions did not reach significance in either the static or the dynamic task, we think that the significant difference observed between the two conditions is not informative with regard to the potential differences in the visual strategies underlying the recognition of static and dynamic facial expression. In fact, the features used with both kinds of stimuli are pretty similar, with the most useful area—the mouth—being the same in both cases.

In addition to comparing the overlap between the significant pixels in the static and dynamic classification images, we calculated a measure of the spread of significant pixels across the face. More specifically, we tested whether participants used facial areas far from the center of the face as much with dynamic as with static stimuli. This could explain the fixation patterns revealed in Experiment 1, showing that participants fixated less on the eyes and mouth and more on the center of the face in the dynamic than in the static condition. The spread was measured by calculating, across all orientations of the face, the average radius of the areas significantly useful. A bootstrap analysis in which we re-created a thousand classification images using random samples (with replacement) of the participants was conducted to get a distribution of the spreads of significant pixels and thus to calculate confidence intervals (CIs) of the difference of spreads of significant pixels in the static and in the dynamic conditions. We found no significant difference for anger (95% CI [−7.1, 68.6]), disgust (95% CI [−19.7, 54.9]), fear (95% CI [−52.5, 109.9]), happiness (95% CI [−72.2, 67.4]), neutrality (95% CI [−37.7, 48.5]), sadness (95% CI [−87.3, 80.0]), or surprise (95% CI [−7.8, 89.1]).

As explained earlier in the description of the Bubbles method, we gathered information about not only which facial areas are most useful to categorize static and dynamic facial expressions but also which spatial frequencies are most useful in these conditions. Thus, we conducted a bootstrap analysis in which we re-created a thousand classification images using random samples (with replacement) of the participants. The classification images were collapsed across expressions to retain only the spatial frequency and space (x, y locations) dimensions. This was done separately for the static and dynamic conditions. The facial areas reaching statistical significance in each spatial frequency band of these classification images were found using a Pixel Test (ts = 4.38, 4.04, 3.69, 3.33, and 3.0 from highest to lowest spatial frequency band). The proportion of significant pixels in each spatial frequency band were measured, and a log-parabola function was fitted on these proportions to estimate the peak of the spatial frequency tuning and the width of the tuning function (Chung, Legge, & Tjan, 2002). The bootstrap gave us a distribution of peaks and widths; we calculated confidence intervals of the difference between the peaks of the spatial frequency tuning function for static and dynamic expressions, as well as confidence intervals of the difference between the widths of the spatial frequency tuning function for static and dynamic expressions. We found no significant difference between the spatial frequency tunings of static and dynamic expressions on either the peaks (95% CI [−10.7, 19.5] cycles per face) or the width (95% CI [−3.4, 3.6]). We also compared the number of significant pixels for static and dynamic classification images separately for each spatial frequency band but found no significant difference for Band 1 (95% CI [−.07, .01]), Band 2 (95% CI [−.85, .26]), Band 3 (95% CI [−.51, .68]), Band 4 (95% [CI −.70, .65]), or Band 5 (95% CI [−.60, .44]).

Taken together, the findings of Experiments 1 and 2 suggest the following: With both kinds of stimuli, participants spent most of the stimulus duration fixating the center of the face. However, with static faces they eventually departed from the center to fixate the eyes and mouth, so they spent more time fixating these features with static than with dynamic expressions. Thus, with static expressions, the results from Experiments 1 and 2 were quite congruent, in that fixating the eyes and mouth was linked with using the eyes and mouth during facial expression recognition. Furthermore, the static Bubbles findings (e.g., Blais et al., 2012; M. L. Smith et al., 2005) are congruent with those obtained using other methods, such as comparing the performance with isolated parts of the face (e.g., Bassili, 1979; Calder, Young, Keane, & Dean, 2000; Dunlap, 1927; Hanawalt, 1944) or using reverse correlation (Jack, Caldara, & Schyns, 2012; Yu et al., 2012). The results obtained with the dynamic expressions were more surprising. Even if observers fixated less on the eyes and mouth, they used these features at least as much as, and sometimes even more than, with static faces. One could argue that applying bubbles on the dynamic stimuli may have modified the fixation pattern and that if we had measured the fixation maps under bubbles constraints, the density of fixations on the features would have been more comparable to the that observed with static faces. Although this argument is not supported by the high similarity between the patterns of hit rates across the different dynamic expressions in Experiments 1 and 2 (r = .92), it cannot be brushed aside.

A third experiment, in which eye fixations were recorded during the recognition of bubblized facial expressions, was thus conducted. The aim of Experiment 3 was to test whether the application of bubbles on the dynamic stimuli changes the fixation pattern such that the density of fixations on features becomes more comparable to that observed with static faces. To do so, we compared the pattern of fixations obtained with bubblized dynamic expressions to those obtained in Experiment 1, without bubbles.

Experiment 3

Method

Participants. Twenty Caucasian participants (six male; mean age = 24 years) took part in the experiment. Note that the aim of the present experiment was to verify only the fixation pattern when bubbles are applied on the stimuli; we did not want to verify the information utilization (i.e., create classification images using the Bubbles). The same number of participants as in Experiment 1 was therefore sufficient. All participants had normal or corrected-to-normal visual acuity. All procedures were carried out with the ethics approval of the Université du Québec en Outaouais.

Materials and stimuli. Stimuli were displayed on a calibrated high-resolution liquid crystal display monitor with a refresh rate of 100 Hz. The experimental program was written in Matlab, using functions from the Psychophysics and EyeLink toolboxes. Eye movements and fixations were measured and recorded at 1000 Hz with the oculomotor system EyeLink 1000. Only the dominant eye was tracked, but viewing was binocular.

The stimuli were the dynamic expressions used in Experiments 1 and 2, but they were presented through dynamic bubbles with the same properties as in Experiment 2. The number of bubbles in the dynamic-bubblized condition was set to that used on average by participants of Experiment 2 to maintain accuracy near the 56.25% target—halfway between floor (12.5% correct) and ceiling (100% correct) accuracy.

Procedure. The procedure was identical to that of Experiment 1, except that there was only one condition, that is, dynamic expressions on which dynamic bubbles were applied.

Results and Discussion

On average, participants correctly categorized the dynamic-bubblized expressions on 70.1% (SD = 20.9) of the trials. The average accuracy rate was of 71.2% (SD = 14.5) with anger, 53.8% (SD = 26.7) with disgust, 66.8% (SD = 13.8) with fear, 93.0% (SD = 7.7) with happiness, 52.8% (SD = 19.0) with neutrality, 74% (SD = 14.5) with sadness, and 78.5% (SD = 14.3) with surprise. The correlation of the accuracy rates obtained in the present experiment and those obtained in Experiment 1 with unaltered dynamic expressions was of r = .75.

On average, participants made 1.32 fixations (SD = .30) per trial, with an average fixation duration of 408.8 ms (SD = 44.0). Participants made significantly fewer fixations, $t(38) = 8.67$, $p < .001$, and longer fixations, $t(38) = 13.09$, $p < .001$, in the dynamic-bubblized condition than in the static condition. Moreover, participants made significantly fewer fixations, $t(38) = 7.34$, $p < .001$, and longer fixations, $t(38) = 9.39$, $p < .001$, in the dynamic-bubblized condition than in the dynamic-unaltered condition. An analysis on the duration of the first fixation was also performed. The average duration of the first fixation was of 421.1 ms (SD = 36.9) in the dynamic-bubblized condition, 258.0 ms (SD = 65.0) in the dynamic-unaltered condition, and 226.6 ms (SD = 55.0) in the static condition. The first fixation was significantly longer in the dynamic-bubblized condition than in the static, $t(38) = 13.13$, $p < .001$, and dynamic-unaltered, $t(38) = 9.75$, $p < .001$, conditions. There was no significant difference on the duration of the first fixation in the static and dynamic-unaltered conditions, $t(38) = 1.64$, $p = .11$.

Using *i*Map, we compared the fixation maps obtained with dynamic-bubblized stimuli to those obtained in Experiment 1, separately for the static and the dynamic conditions. Two 7 (facial expressions: anger, disgust, fear, happiness, neutrality, sadness, surprise) × 2 (dynamic bubblized vs. static or dynamic conditions from Experiment 1) mixed pixelwise ANOVAs, followed by bootstraps, were conducted. We found a significant main effect of task when we compared the dynamic-bubblized condition of Experiment 3 with the static-unaltered condition of Experiment 1, but neither the main effect of emotion nor the interaction between emotion and task reached statistical threshold. This replicates the results obtained by contrasting the static-unaltered and the dynamic-unaltered conditions of Experiment 1. The results, presented in Figure 6, show a significantly higher density of fixation in the eye area in the static than in the dynamic-bubblized condition, which replicates the finding of Experiment 1 when the pattern of fixations obtained with static and dynamic unaltered stimuli were compared. However, contrary to the case in Experiment 1, the density of fixations on the mouth area and on the center of the faces did not differ significantly in the two conditions. Most important, there was no significant difference between the dynamic-bubblized and the dynamic-unaltered conditions.

As explained earlier, the aim of this third experiment was to test whether bubbles modify the strategy such that the features would be as fixated as they are with static stimuli, a possible explanation for the similarity of information utilization found in Experiment 2. The results indicate that the finding of Experiment 2 cannot be explained by this hypothesis. In fact, even when dynamic bubbles were applied on facial expressions, the participants fixated the eye area significantly less than with static-unaltered expressions, and the fixation pattern did not differ significantly from the one obtained with dynamic-unaltered expressions. Moreover, an analysis

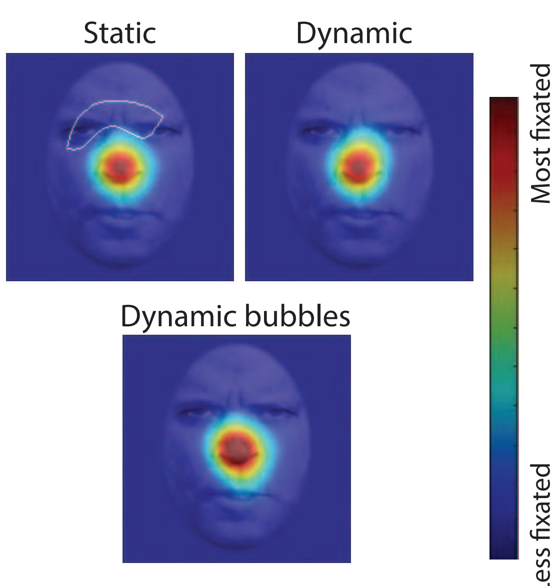

Figure 6. Maps of fixations obtained with unaltered static and dynamic stimuli in Experiment 1 (upper row) and map of fixation obtained with dynamic-bubblized stimuli in Experiment 3 (lower row). The area that was fixated significantly more in the static than in the dynamic-bubblized condition is outlined in white.

on the number of fixations and fixation duration indicated that participants' gaze remained for a longer duration on the location it was on when the face first appeared, which further argues against the hypothesis that applying bubbles on a face may drive participants to rely on a strategy where they fixate more directly the features compared to when there is no bubble.

General Discussion

The main aim of the present study was to compare the ocular fixation patterns for static and dynamic facial expressions. This is, to the best of our knowledge, the first study to directly compare the visual strategies used to recognize static and dynamic facial expressions displayed by the same set of actors.

The results of Experiment 1 showed that the pattern of fixations did not differ as a function of the expression presented. This finding was replicated in Experiment 3 and is congruent with results of previous studies that did not find an effect of emotion on the fixation patterns with static expressions (Jack et al., 2009; Vaidya et al., 2014). Here, we showed for the first time that this was also true with dynamic expressions. Furthermore, we discovered that the fixation pattern with dynamic expressions is different from what was previously observed with static expressions: Fewer fixations were made on the main facial features (i.e., eyes and mouth) with dynamic compared with static expressions.

The different fixation patterns were observed using short stimulus duration (500 ms), thereby limiting the number of fixations included in the analysis. The stimulus duration was chosen based on the longest natural duration of the facial expressions performed by the actors who were filmed for the STOIC database. In the present study, we therefore measured the eye fixations occurring during the natural duration of a simulated dynamic expression. It is quite possible that the pattern of fixations observed with static and dynamic expressions would have been different if longer stimulus duration had been used. Nevertheless, we think that using the natural duration of expressions provides a more realistic comparison of the eye fixations underlying the recognition of static and dynamic facial expressions. Moreover, studies on face identification have shown that two fixations suffice for the recognition of faces (Hsiao & Cottrell, 2008; Or, Peterson, & Eckstein, 2015). Although no studies have yet examined the number of fixations sufficient for recognizing facial expressions, the number of fixations made by our participants while recognizing facial expressions and their high hit rates suggest a similar figure.

The results of Experiment 2 showed that, except for the surprise expression, the information used with static expressions was similar to that used with dynamic expressions and that when differences were revealed they were mostly quantitative. For instance, the eyes and wrinkles areas were utilized to recognize both static and dynamic anger expressions, but they were utilized more in the dynamic than in the static condition. The surprise expression was an exception to this pattern of results: The eyes area was not significantly utilized with the static surprise expression, whereas it was significantly utilized with the dynamic surprise expression. Thus, our results are congruent with those obtained by Nusseck et al. (2008), who showed a large overlap between the areas containing diagnostic motion for the categorization of dynamic expressions, and the diagnostic facial areas for the categorization of static expressions revealed by M. L. Smith et al. (2005). Our results are also in agreement with those of Yu et al. (2012), who showed a high correlation between the action units coded in visual memory and the action units actually available in facial expressions. In both these studies, however, some differences were revealed between the facial features or action units used with static and dynamic expressions, and these differences were difficult to interpret because the results with static and dynamic expressions were obtained in separate studies using different methodologies and different sets of stimuli. Here, static and dynamic facial expressions were directly compared using the same set of facial expressions and the same methodology. The results confirm that the facial features underlying facial expression recognition are similar for static and dynamic stimuli.

The results also provide more information regarding the specific pattern of feature utilization for each emotion. If one compares the pattern across expressions, one noticeable difference is that more facial areas were useful for the recognition of anger than for the recognition of the other emotions. This suggests that both the eye or frown lines and the mouth convey information about that emotion, whereas for the other emotions, the diagnostic information was typically limited to a single area. Note, however, that in the results presented in Figure 5, the spatial frequency dimension was collapsed. One consequence of this is that the features that are useful in only one spatial frequency band may not reach the statistical threshold, whereas those that are useful in many spatial frequency bands will tend to reach the threshold. An analysis of the facial areas used in each spatial frequency band was also conducted (see Figure S1 in the online supplemental materials). From this analysis, one can see that the eyes and mouth areas were useful in at least three of the five spatial frequency bands for the anger expression, whereas for the other expressions they were never used in more than two spatial frequency bands. The diagnosticity of more spatial frequency bands for anger than for other expressions may perhaps be linked to the evolutionary importance of being able to efficiently recognize that expression, whether the individual displaying it is at close distance or farther away (F. W. Smith & Schyns, 2009). Such results are also consistent with those of studies showing that an anger expression can automatically capture visual attention (Fox et al., 2000; Huang, Chang, & Chen, 2011; Öhman, Soares, Juth, Lindström, & Esteves, 2012; see, however, Becker, Anderson, Mortensen, Neufeld, & Neel, 2011). It is also interesting to note (in Figure S1 of the online supplemental materials) that the eyes area is significantly useful in high spatial frequencies for the recognition of fear, which closely replicates the results of previous studies (Adolphs et al., 2005; M. L. Smith et al., 2005).

The specific emotional categories included in the task have likely modulated the visual strategies revealed. In fact, in a recent study, M. L. Smith and Merlusca (2014) showed that the information used to successfully categorize an expression changes as a function of the specific emotion labels included in a categorization task. Likewise, the number of identities and the specific way in which each of these identities express the emotions will also modulate the diagnostic information for the recognition of each expression. The pattern of eye fixations is also likely to be in part attributable to the specific characteristics of the task in which the participants took part. For instance, as mentioned in the introduction, the finding of similar versus different eye fixation patterns as a function of the emotion appears to depend on the task used: The categorization of the six basic expressions at their apex most often lead to fixation patterns that do not change as a function of the expression presented, whereas other tasks like intensity judgments, deciding whether the expression presented is part of one predefined category, or categorizing subtle expressions have revealed different patterns for different emotions.

Altogether, the results of the three experiments presented here suggest that the differences observed in the fixation patterns are not linked to a differential use of the facial features during the recognition of facial expressions. It is important to note, however, that the dynamic Bubbles method used in Experiment 2 did not examine the use of motion information. Thus, similar facial areas may be revealed in static and dynamic classification images but for different reasons. For example, the finding that participants used the mouth area to recognize dynamic happy expressions may indicate that they processed its shape, its motion, or both. It has been shown that biological motion (e.g., the motion occurring in facial expressions) can be processed outside the fovea (Gurnsey, Roddy, Ouhnana, & Troje, 2008; Thompson, Hansen, Hess, & Troje, 2007). Thus, fixating mostly the center of dynamic facial expressions may have allowed the participants to extract biological motion information occurring around the main facial areas. This would predict a higher utilization of the low spatial frequencies with dynamic than with static stimuli. But we found no such difference in the peak or in the bandwidth of the spatial frequency tuning functions in Experiment 2. However, the present study was not designed in the best way to compare spatial frequency tuning for both kinds of stimuli. Although in the static condition one bubble appeared for the whole duration of a stimulus, in the dynamic condition a bubble's duration varied as a function of the spatial frequency band it was sampling, with increasing duration as the spatial frequency band increased. This manipulation was done to take into account the faster processing of lower spatial frequencies (e.g., Hughes et al., 1990; Parker et al., 1992). Nevertheless, it makes the comparison of spatial frequency utilization between static and dynamic expressions more difficult, so more research will be necessary to assess this hypothesis.

Another hypothesis to explain the finding of similar facial features utilization despite different fixation patterns for static and dynamic expressions is that motion may make features more salient and therefore attract attention toward them (see also Tobin, Favelle, & Palermo, 2016). A similar proposition has been made by Horstmann and Ansorge (2009); they suggested that the advantage observed with dynamic expressions is in part due to motion's automatically capturing attention to dynamic stimuli. Hence, it has been shown that attention increases spatial resolution (Yeshurun & Carrasco, 1998). Directing attention toward the moving eyes and mouth could therefore increase the spatial resolution with which these areas are processed and reduce the need to directly fixate them.

On the other hand, one could propose that the finding of different ocular patterns for static and dynamic expressions reflects that individuals are less likely to look away from their initial fixation location when other potentially informative areas of the face are moving. In that regard, the analysis on the duration of the first fixation is particularly informative. In fact, the duration of the first fixation did not significantly differ between the static and dynamic-unaltered conditions. This suggests that participants made a saccade following a similar delay after stimulus onset in both conditions. However, in the dynamic-unaltered condition, the next fixation was again directed close to the center of the face. Thus, the results do not support the hypothesis that one is simply less likely to look away from its initial fixation location when there is motion in the stimulus; they rather suggest that with static and dynamic stimuli, the fixations are directed toward different locations during the stimulus processing.

More research is needed to investigate the reason underlying the use of different fixation patterns with static and dynamic expressions. In that regard, the two hypotheses proposed earlier—that is, that participants fixate more on the center of the face with dynamic stimuli either because they process biological motion in periphery or because motion makes the features more salient and attracts attention toward them, thereby increasing their spatial resolution and decreasing the need to directly fixate the feature—need further investigation. These hypotheses have one point in common: They imply that fixations are directed toward the positions that will optimize information processing given the stimulus constraints, a proposition that has been supported by many studies in face processing (Hsiao & Cottrell, 2008; Peterson & Eckstein, 2012; Zerouali, Lina, & Jemel, 2013). Following this line of thought, the present results predict that the sequence of fixations that will optimize facial expression processing should differ for static and dynamic expressions. Moreover, such a hypothesis may explain the finding of a longer first fixation when dynamic bubbles are applied over the stimulus. In fact, applying bubbles on the stimuli probably made more unpredictable the next optimal position to fixate, because a bubble is likely to have disappeared before the saccade lands on the new area. Relatedly, studying the time course of information utilization with static facial expressions could further the understanding of the mechanisms underlying the different eye fixation patterns with static and dynamic expressions.

Although the utilization of dynamic stimuli in the present study allowed us to take a humble step toward understanding the recognition of emotions in a more ecological setting, it is important to keep in mind that the stimuli used in the present study consisted of simulated rather than spontaneous facial expressions. Simulated expressions are used in most studies attempting to understand the mechanisms underlying the recognition of facial expressions. However, spontaneous and simulated facial expressions differ in laterality of motion at onset (Ross & Pulusu, 2013) and in the time course of the facial action units involved (Cohn & Schmidt, 2004). Thus, it would be interesting in future research to examine the visual strategies underlying the recognition of static and dynamic spontaneous expressions of emotion.

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,* 68–72. http://dx.doi.org/10.1038/nature03086

Ahumada, A. J., Jr., & Lovell, J. (1971). Stimulus features in signal detection. *Journal of the Acoustical Society of America, 49,* 1751–1756. http://dx.doi.org/10.1121/1.1912577

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,* 403–410. http://dx.doi.org/10.1111/j.0956-7976.2005.01548.x

Bassili, J. N. (1979). Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face. *Journal of Personality and Social Psychology, 37,* 2049–2058. http://dx.doi.org/10.1037/0022-3514.37.11.2049

Beaudry, O., Roy-Charland, A., Perron, M., Cormier, I., & Tapp, R. (2014). Featural processing in recognition of emotional facial expressions. *Cognition and Emotion, 28,* 416–432. http://dx.doi.org/10.1080/02699931.2013.833500

Becker, D. V., Anderson, U. S., Mortensen, C. R., Neufeld, S. L., & Neel, R. (2011). The face in the crowd effect unconfounded: Happy faces, not angry faces, are more efficiently detected in single- and multiple-target visual search tasks. *Journal of Experimental Psychology: General, 140,* 637–659. http://dx.doi.org/10.1037/a0024060

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PLoS ONE, 3*(8), e3022. http://dx.doi.org/10.1371/journal.pone.0003022

Blais, C., Roy, C., Fiset, D., Arguin, M., & Gosselin, F. (2012). The eyes are not the window to basic emotions. *Neuropsychologia, 50,* 2830–2838. http://dx.doi.org/10.1016/j.neuropsychologia.2012.08.010

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436. http://dx.doi.org/10.1163/156856897X00357

Buchan, J. N., Paré, M., & Munhall, K. G. (2007). Spatial statistics of gaze fixations during dynamic face processing. *Social Neuroscience, 2,* 1–13. http://dx.doi.org/10.1080/17470910601043644

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 527–551. http://dx.doi.org/10.1037/0096-1523.26.2.527

Chauvin, A., Worsley, K. J., Schyns, P. G., Arguin, M., & Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *Journal of Vision, 5,* 1. http://dx.doi.org/10.1167/5.9.1

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. http://dx.doi.org/10.1024/1421-0185/a000061

Chung, S. T., Legge, G. E., & Tjan, B. S. (2002). Spatial-frequency characteristics of letter identification in central and peripheral vision. *Vision Research, 42,* 2137–2152. http://dx.doi.org/10.1016/S0042-6989(02)00092-5

Cohn, J. F., & Schmidt, K. L. (2004). The timing of facial motion in posed and spontaneous smiles. *International Journal of Wavelets, Multiresolution, and Information Processing, 2,* 121–132. http://dx.doi.org/10.1142/S021969130400041X

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments & Computers, 34,* 613–617. http://dx.doi.org/10.3758/BF03195489

Coutrot, A., & Guyader, N. (2014). How saliency, faces, and sound influence gaze in dynamic social scenes. *Journal of Vision, 14,* 5. http://dx.doi.org/10.1167/14.8.5

Cunningham, D. W., & Wallraven, C. (2009). Dynamic information for the recognition of conversational expressions. *Journal of Vision, 9,* 7. http://dx.doi.org/10.1167/9.13.7

Dunlap, K. (1927). The role of eye-muscles and mouth-muscles in the expression of the emotions. *Genetic Psychology Monographs, 2,* 196–233.

Dupuis-Roy, N., Dufresne, K., Fiset, D., & Gosselin, F. (2012). The time course of chromatic and achromatic information extraction in a face-gender discrimination task. *Perception ECVP Abstract, 41,* 114.

Dupuis-Roy, N., Fortin, I., Fiset, D., & Gosselin, F. (2009). Uncovering gender discrimination cues in a realistic setting. *Journal of Vision, 9,* 10. http://dx.doi.org/10.1167/9.2.10

Eisenbarth, H., & Alpers, G. W. (2011). Happy mouth and sad eyes: Scanning emotional facial expressions. *Emotion, 11,* 860–865. http://dx.doi.org/10.1037/a0022758

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system: A technique for the measurement of facial movement*. Palo Alto, CA: Consulting Psychologists Press.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). Thousand Oaks, CA: Sage.

Fox, E., Lester, V., Russo, R., Bowles, R. J., Pichler, A., & Dutton, K. (2000). Facial expressions of emotion: Are angry faces detected more efficiently? *Cognition and Emotion, 14,* 61–92. http://dx.doi.org/10.1080/026999300378996

Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research, 41,* 2261–2271. http://dx.doi.org/10.1016/S0042-6989(01)00097-9

Gurnsey, R., Roddy, G., Ouhnana, M., & Troje, N. F. (2008). Stimulus magnification equates identification and discrimination of biological motion across the visual field. *Vision Research, 48,* 2827–2834. http://dx.doi.org/10.1016/j.visres.2008.09.016

Hammal, Z., Gosselin, F., & Fortin, I. (2009). How efficient are the recognition of dynamic and static facial expressions? *Journal of Vision, 9,* 499. http://dx.doi.org/10.1167/9.8.499

Hanawalt, N. G. (1944). The role of the upper and the lower parts of the face as a basis for judging facial expressions: II. In posed expressions and "candid-camera" pictures. *Journal of General Psychology, 31,* 23–36. http://dx.doi.org/10.1080/00221309.1944.10545217

Horstmann, G., & Ansorge, U. (2009). Visual search for facial expressions of emotions: A comparison of dynamic and static faces. *Emotion, 9,* 29–38. http://dx.doi.org/10.1037/a0014147

Hsiao, J. H. W., & Cottrell, G. (2008). Two fixations suffice in face recognition. *Psychological Science, 19,* 998–1006. http://dx.doi.org/10.1111/j.1467-9280.2008.02191.x

Huang, S. L., Chang, Y. C., & Chen, Y. J. (2011). Task-irrelevant angry faces capture attention in visual search while modulated by resources. *Emotion, 11,* 544–552. http://dx.doi.org/10.1037/a0022763

Hughes, H. C., Fendrich, R., & Reuter-Lorenz, P. A. (1990). Global versus local processing in the absence of low spatial frequencies. *Journal of Cognitive Neuroscience, 2,* 272–282. http://dx.doi.org/10.1162/jocn.1990.2.3.272

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,* 173–181. http://dx.doi.org/10.1016/0028-3932(93)90045-2

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

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

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,* 187–192. http://dx.doi.org/10.1016/j.cub.2013.11.064

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

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

Kilts, C. D., Egan, G., Gideon, D. A., Ely, T. D., & Hoffman, J. M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *NeuroImage, 18,* 156–168. http://dx.doi.org/10.1006/nimg.2002.1323

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception, 36,* 1.

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,* 1023–1033. http://dx.doi.org/10.1093/cercor/13.10.1023

Lao, J., Miellet, S., Pernet, C., Sokhn, N., & Caldara, R. (2016). *i*Map4*i*Map: An open source toolbox for the statistical fixation mapping of eye movement data with linear mixed modeling. *Behavior Research Methods*. Advance online publication. http://dx.doi.org/10.3758/s13428-016-0737-x

Lischke, A., Berger, C., Prehn, K., Heinrichs, M., Herpertz, S. C., & Domes, G. (2012). Intranasal oxytocin enhances emotion recognition from dynamic facial expressions and leaves eye-gaze unaffected. *Psychoneuroendocrinology, 37,* 475–481. http://dx.doi.org/10.1016/j.psyneuen.2011.07.015

MATLAB and Statistics Toolbox Release. (2006). Natick, Massachusetts: The MathWorks, Inc.

Matsuzaki, N., & Sato, T. (2008). The perception of facial expressions from two-frame apparent motion. *Perception, 37,* 1560–1568. http://dx.doi.org/10.1068/p5769

Nusseck, M., Cunningham, D. W., Wallraven, C., & Bülthoff, H. H. (2008). The contribution of different facial regions to the recognition of conversational expressions. *Journal of Vision, 8,* 1. http://dx.doi.org/10.1167/8.8.1

Öhman, A., Soares, S. C., Juth, P., Lindström, B., & Esteves, F. (2012). Evolutionary derived modulations of attention to two common fear stimuli: Serpents and hostile humans. *Journal of Cognitive Psychology, 24,* 17–32. http://dx.doi.org/10.1080/20445911.2011.629603

Or, C. C. F., Peterson, M. F., & Eckstein, M. P. (2015). Initial eye movements during face identification are optimal and similar across cultures. *Journal of Vision, 15,* 12. http://dx.doi.org/10.1167/15.13.12

Parker, D. M., Lishman, J. R., & Hughes, J. (1992). Temporal integration of spatially filtered visual images. *Perception, 21,* 147–160. http://dx.doi.org/10.1068/p210147

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442. http://dx.doi.org/10.1163/156856897X00366

Peterson, M. F., & Eckstein, M. P. (2012). Looking just below the eyes is optimal across face recognition tasks. *Proceedings of the National Academy of Sciences of the United States of America, 109*(48), E3314–E3323. http://dx.doi.org/10.1073/pnas.1214269109

Ross, E. D., & Pulusu, V. K. (2013). Posed versus spontaneous facial expressions are modulated by opposite cerebral hemispheres. *Cortex, 49,* 1280–1291. http://dx.doi.org/10.1016/j.cortex.2012.05.002

Roy, C., Blais, C., Fiset, D., Rainville, P., & Gosselin, F. (2015). Efficient information for recognizing pain in facial expressions. *European Journal of Pain, 19,* 852–860. http://dx.doi.org/10.1002/ejp.676

Roy, S., Roy, C., Fortin, I., Ethier-Majcher, C., Belin, P., & Gosselin, F. (2007). A dynamic facial expression database. *Journal of Vision, 7,* 944. http://dx.doi.org/10.1167/7.9.944

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. *Cognitive Brain Research, 20,* 81–91. http://dx.doi.org/10.1016/j.cogbrainres.2004.01.008

Schultz, J., & Pilz, K. S. (2009). Natural facial motion enhances cortical responses to faces. *Experimental Brain Research, 194,* 465–475. http://dx.doi.org/10.1007/s00221-009-1721-9

Schurgin, M. W., Nelson, J., Iida, S., Ohira, H., Chiao, J. Y., & Franconeri, S. L. (2014). Eye movements during emotion recognition in faces. *Journal of Vision, 14,* 1. http://dx.doi.org/10.1167/14.13.14

Smith, F. W., & Schyns, P. G. (2009). Smile through your fear and sadness: Transmitting and identifying facial expression signals over a range of viewing distances. *Psychological Science, 20,* 1202–1208. http://dx.doi.org/10.1111/j.1467-9280.2009.02427.x

Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and decoding facial expressions. *Psychological Science, 16,* 184–189. http://dx.doi.org/10.1111/j.0956-7976.2005.00801.x

Smith, M. L., & Merlusca, C. (2014). How task shapes the use of information during facial expression categorizations. *Emotion, 14,* 478–487. http://dx.doi.org/10.1037/a0035588

Thompson, B., Hansen, B. C., Hess, R. F., & Troje, N. F. (2007). Peripheral vision: Good for biological motion, bad for signal noise segregation? *Journal of Vision, 7,* 12. http://dx.doi.org/10.1167/7.10.12

Tobin, A., Favelle, S., & Palermo, R. (2016). Dynamic facial expressions are processed holistically, but not more holistically than static facial expressions. *Cognition and Emotion, 30,* 1208–1221. http://dx.doi.org/10.1080/02699931.2015.1049936

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. http://dx.doi.org/10.1016/j.brainres.2009.05.075

Vaidya, A. R., Jin, C., & Fellows, L. K. (2014). Eye spy: The predictive value of fixation patterns in detecting subtle and extreme emotions from faces. *Cognition, 133,* 443–456. http://dx.doi.org/10.1016/j.cognition.2014.07.004

Vinette, C., Gosselin, F., & Schyns, P. G. (2004). Spatio-temporal dynamics of face recognition in a flash: It's in the eyes! *Cognitive Science, 28,* 289–301.

Võ, M. L. H., Smith, T. J., Mital, P. K., & Henderson, J. M. (2012). Do the eyes really have it? Dynamic allocation of attention when viewing moving faces. *Journal of Vision, 12,* 3. http://dx.doi.org/10.1167/12.13.3

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

Yeshurun, Y., & Carrasco, M. (1998, November 5). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature, 396,* 72–75. http://dx.doi.org/10.1038/23936

Yu, H., Garrod, O., Jack, R., & Schyns, P. (2015). A framework for automatic and perceptually valid facial expression generation. *Multimedia Tools and Applications, 74,* 9427–9447. http://dx.doi.org/10.1007/s11042-014-2125-9

Yu, H., Garrod, O. G., & Schyns, P. G. (2012). Perception-driven facial expression synthesis. *Computers & Graphics, 36,* 152–162. http://dx.doi.org/10.1016/j.cag.2011.12.002

Zerouali, Y., Lina, J. M., & Jemel, B. (2013). Optimal eye-gaze fixation position for face-related neural responses. *PLoS ONE, 8*(6), e60128. http://dx.doi.org/10.1371/journal.pone.0060128

Received October 5, 2015
Revision received December 20, 2016
Accepted January 5, 2017 ■