A PSYCHOPHYSICAL EXPLORATION OF THE PERCEPTION OF EMOTION FROM ABSTRACT ART*

STÉPHANIE DUBAL
Hôpital Pitié-Salpêtrière, Paris, France

JÉRÔME PELLETIER
Institut Jean Nicod, Paris, France

AURE-ELISE LEREBOURS
Hôpital Pitié-Salpêtrière, Paris, France

YOLAINE ESCANDE
Centre de Recherches sur les Arts et le Langage, Paris, France

MARINE TAFFOU
Hôpital Pitié-Salpêtrière, Paris, France

KENNETH KNOBLAUCH
Stem-Cell and Brain Research Institute, Lyon, France
and Université Lyon, Lyon, France

ABSTRACT

Centuries ago, western philosophers of art proposed that an artifact may express emotion. The much older tradition of Chinese calligraphy and landscape painting sees in the brushstroke a central element of the emotional expressiveness of pictorial art. How emotion is detected and perceived from artwork is a current subject of exploration in psychology. We used Signal Detection Theory to determine whether or not naive subjects were able to detect the emotional classification proposed by an artist. Thirty pairs of

*This work was funded by a grant from the ANR (French National Research Agency; 10-CREA-005).

© 2014, Baywood Publishing Co., Inc.
doi: http://dx.org/10.2190/EM.32.1..EOV.4
http://baywood.com
Chinese characters were produced by an artist so as to render the same Chinese sign with different emotional expressive intensities: one with low emotional expressivity and a second with high emotional expressivity. Twenty-two participants were asked to rate emotional intensity on a 5-point scale while viewing each calligraphy. $d'$, a sensitivity measure reflecting the capacity to detect a signal from a noisy background, was estimated from ROC curves. If the participants categorized the stimuli according to the artist’s classification (or in direct opposition), $d'$ would differ significantly from 0. Sixteen of the participants had $d'$ values significantly different from 0, thereby showing that they discriminated the emotional categorization of the calligraphy. Those results indicate a significant sensitivity to the emotion expressed in a Chinese calligraphy by a non-Chinese speaking population ignorant of the art of calligraphy and the artist’s intention.

Pictorial art generates a rich fund of experiences, which involve a significant part of our mental life, from perception to cognition, from creation to reception. The emotional side of our relation with pictorial works remains poorly understood. In the West, since the end of the 18th century and especially the 19th century, the expression theory of art has claimed that an artifact expresses emotions (Collingwood, 1945; Tolstoy, 1962). The much older tradition of Chinese calligraphy and landscape painting sees in the brushstroke a central element of the emotional expressiveness of pictorial art (Cheng, 1986; Escande, 2005). The theoretical tradition of Chinese visual art can be traced back to the first centuries AD; since its beginning, the expression of emotions was fundamental to its appreciation. This appreciation is based on the use of the brush, with black ink, on a sheet of paper, for writing as well as for painting. The brushstroke itself ($hua$) serves as a definition both of the act of writing and painting, and of its result. In this tradition, the value of an artwork is generally believed to be related to its emotional expressivity.

Therefore, for the Chinese and for readers of Chinese writing, calligraphic art is never an abstract art since the characters themselves have meaning. However, for naive non-Chinese subjects, the brushstrokes are semantically meaningless, and thus they can be treated as abstract art. Nevertheless, the perception of emotion based on abstract art has not yet been studied in the field of Chinese visual arts. This is because although there is a rich theorization in Chinese, in all cases attention is focussed on the expressivity of the gesture, on the creative movement (Cheng, 1979; Shen, 1977), rather than the perception of the viewer. For example, Hsiung Ping-Ming’s *theoretical Systems of Chinese Calligraphy* (1984) devotes one chapter to the question of “expression of emotions” in calligraphy and to its appreciation and aesthetic value. Chinese theoretical tradition considers that words are insufficient to transmit emotions, and that the brushstroke is the vehicle of non-semantic expression. But in his book, Hsiung’s viewpoint is mainly focused on the creative process, and not on the viewer’s perception.
In the field of psychology, a number of studies have focused on the properties of stimuli that convey emotion, in artistic stimuli as well as in natural stimuli, such as faces (Adolphs, 2002; Bar & Neta, 2006; Reber, Schwarz, & Winkielman, 2004). This is especially clear in the example of a happy face in which a very characteristic pattern can transmit emotion. The region of the mouth provides the most important diagnostic feature for recognition or identification of happy facial expressions (Adolphs, 2002; Leppanen & Hietanen, 2007; Schyns, Petro, & Smith, 2007), and even an upward lip pattern is sufficient for a face to be classified as happy (Dubal, Foucher, Jouvent, & Nadel, 2011). In artistic stimuli, including abstract art, it is less known what the physical properties are that contribute to the expression of affect. For Reber et al. (2004), aesthetic pleasure is a function of the perceiver’s processing dynamics; the more fluently perceivers can process an object, the more positive their aesthetic response. Physical properties of artistic stimuli influence processing fluency; color, contrast, and symmetry are factors that contribute to the modulation of aesthetic appraisal via fluency. For example, the nature of an object’s contour provides a source of influence of preferences (Bar & Neta, 2006). Specifically, objects with pointed features and sharp angles are liked less than objects with curved features. Furthermore, the observation that the interpretability of a visual scene increases the pleasure evoked by the scene supports the fluency hypothesis (Biederman & Vessel, 2006). It appears that at least in part, perceptual features can contribute to the emotional quality of a stimulus.

Brain studies of affective processing also bring elements to the hypothesis that perceptual features can contribute to affect. One line of evidence in support of this argument comes from studies that show higher activity in visual cortical regions in response to emotional than neutral stimuli, including art stimuli (e.g., Mourão-Miranda, Volchan, Moll, Oliveira-Souza, Oliveira, Bramati, et al., 2003; Vartanian & Goel, 2004), but also more generally emotional scenes and expressive faces (e.g., Britton, Taylor, Sudheimer, & Liberzon, 2006; Junghofer, Sabatinelli, Bradley, Schupp, Elbert, & Lang, 2006; Lane, Chua, & Dolan, 1999; Lane, Fink, Chau, & Dolan, 1997; Lang, Bradley, Fitzsimmons, Cuthbert, Scott, Moulder, et al., 1998; Mitterschiffthaler, Kumari, Malhi, Brown, Giampietro, Brammer, et al., 2003; Sabatinelli, Bradley, Fitzsimmons, & Lang, 2005; Taylor, Liberzon, Decker, & Koepepe, 2000). A second line of evidence comes from studies using techniques such as EEG and MEG, which allow the measurement of the temporal dynamics of the processing of emotional stimuli. In this domain, some studies have found that emotion from various categories of stimuli is encoded very early, as early as the first 100 ms poststimulus, at the level of perceptual encoding (Chammat, Foucher, Nadel, & Dubal, 2010; Dubal et al., 2011). Indeed, cortical visual responses in this early time window are higher in emotional than neutral categories for human faces (Batty & Taylor, 2003) and schematic emotional faces (Eger, Jedynak, Iwaki, & Skrandies, 2003), but also for emotionally arousing stimuli such as body expressions (Meeren, van Heijnsbergen, & de Gelder, 2005).
words (Ortigue, Michel, Murray, Mohr, Carbonnel, & Landis, 2004) or complex visual scenes (Alorda, Serrano-Pedraza, Campos-Bueno, Sierra-Vazquez, & Montoya, 2007; Carretie, Hinojosa, Martin-Loeches, Mercado, & Tapia, 2004). Such perceptual enhancement has been found in response to positive as well as to negative eliciting stimuli (Batty & Taylor, 2003; Brosch, Sander, Pourtois, & Scherer, 2008).

As pointed out by Keil, Stolarova, Moratti, and Ray (2007), at the level of perceptual encoding affective properties may be confounded with physical properties of the stimulus. In this regard, it has been shown that the initial sensory response in the visual cortex is sensitive to simple features associated with emotionality (Pourtois, Thut, Grave de Peralta, Michel, & Vuilleumier, 2005). Thus, several lines of evidence from philosophy, psychology, and neuroscience converge to show that physical properties can contribute to the transmission of emotional expressiveness (Keil et al., 2007; Leder, Belke, Oeberst, & Augustin, 2004; Reber et al., 2004).

We explored the perception of emotion from Chinese calligraphy by naive subjects, ignorant of the semantic meaning of the ideograms to test the hypothesis that in calligraphy, emotion can be extracted from some of the features of the stimuli independently of the meaning. To do so we measured the perception of emotion from Chinese calligraphy depicting different emotional intensities. The objective of the experiment was to evaluate, using psychophysical methods, the sensitivity of observers to detect differences in the emotional intensity of Chinese calligraphic characters. More specifically, the objective of the experiment was to evaluate firstly whether a non-Chinese sample without any knowledge of the Chinese language and without practice or knowledge of the art of calligraphy would classify calligraphy at the emotional level non-randomly. The second objective was to determine if and to what extent a classification of calligraphy at the emotional level by the same sample would match the artist’s classification of the emotional level of the ideograms created specifically for this experiment.

**METHOD**

**Participants**

Twenty-two right-handed participants volunteered for this study (11 women, mean age: 22.4 ± 2.59). All had normal or corrected-to-normal vision and no neurological disorder. All were naive to Chinese calligraphy, and the Chinese language, in general. All participants provided informed consent to take part in the experiment that had been previously approved by the local ethical committee. Their general level of art experience was ascertained with a screening questionnaire (Chatterjee, Widick, Sternschein, Smith, & Bromberger, 2010).
Material

Thirty Chinese signs were selected and produced by an artist to render two versions of the same Chinese sign expressing different emotional intensities. Thus, for each of the 30 Chinese signs the artist created two instances, one with low emotional expressivity (LEE) and a second with high emotional expressivity (HEE) (see Figure 1 for an example). All stimuli were black and white and were equated for average luminance and size. Images were presented in the center of the computer screen and subtended a vertical visual angle of 10.05° and a horizontal visual angle of 9.41°.

Procedure

Participants were explicitly informed that they would be shown abstract artworks. Each participant was seated in front of a screen and was presented with the 60 calligraphic signs in random order. On each trial, the participant was asked to respond on a scale of 1-5 as to the emotional intensity perceived from the stimulus (i.e., calligraphy), for which 1 was used if he/she felt that the character expressed a low level of intensity and 5 for the high level of intensity. The participant was directed to respond with category 3 if the stimulus was of intermediate emotional intensity.

A block of eight practice trials preceded the 60 experimental trials. Each trial began with the presentation at the center of the screen of a fixation cross for 200 ms, followed by a stimulus for 2000 ms. The inter-trial interval was fixed at 1000 ms.

Figure 1. Example of a pair of low emotional expressivity (LEE) (left) and high emotional expressivity (HEE) (right) signs.
RESULTS

Analysis

The analysis of the data was based on Signal Detection Theory (SDT), which assumes that the perceiver’s capacity for detecting a signal is constrained by internal responses to the stimuli and contamination by noise (MacMillan & Creelman, 2005). The 60 stimuli were divided into two categories, LEE and HEE, according to the artist’s classification. In this context, the signal is the difference of emotional intensity evoked by the LEE and HEE calligraphies, LEE being noise and HEE being noise plus signal (see Figure 2). Our goal was to determine whether or not naive participants were able to detect the signal (i.e., the emotional classification proposed by the artist).

The principal measure that was extracted from the data was the so-called d’ index. d’ is a measure of sensitivity that reflects the capacity of detecting a signal from a noisy background, and corresponds to the estimated difference in the distributions of internal responses evoked by the two classes of stimuli (Figure 2). According to SDT, if the distributions of internal responses for two classes of stimuli overlap, then the observer will choose a criterion value to differentially classify the stimuli. When the internal response is greater than this criterion, the observer chooses one category, otherwise the other. Suppose the criterion is chosen at an internal response of 2, as in Figure 2. Then when the internal response is greater than 2, the observer responds that the stimulus was of HEE. If the stimulus was from the LEE group, then the response is labeled a “Hit.” If the stimulus was from the HEE group, then the response is labeled a “False Alarm.”

With respect to the decision model in Figure 2, the proportion of Hits for the assumed criterion is indicated by the area of the vertical hatched region. The proportion of False Alarms is indicated by the horizontal hatched region. These two measures of performance determine the separation of the two underlying distributions—d’. In the rating scale experiment, we asked the participants to employ several criteria simultaneously, indicated by the four vertical dashed lines that separate the decision space into regions for each of the rating levels. For an ideal observer, any of these criteria would suffice to determine the value of d’. Given observer variability, however, it is better to obtain the estimates by combining all four rating (or criterion) boundaries. This is done by plotting the Hit rate against the False Alarm rate for each criterion used, generating what is called a Receiver Operator Characteristic (ROC) curve. We can then determine what separation of the two underlying distributions best characterizes the set of Hit and False Alarm rates.

Figure 3 shows the Hit rate plotted against the False Alarm rate for three observers that span the results obtained in our sample of participants with the best fitting ROC curves and estimated d’ values indicated in the insets. The data from the left and right participants each show strong evidence for the ability to
Figure 2. Schematic decision space in the rating scale experiment.

Notes: LEE and HEE represent hypothetical distributions of internal response for Low and High Emotional Effects, respectively, with separation here of $d' = 1$. The vertical dashed lines indicate hypothetical criteria for each value of the rating scale. The hatched regions correspond to Hit (vertical) and False Alarm (horizontal) rates for the criterion of ratings greater than or equal to 4.
classify the stimuli along the same dimensions as the artist. The difference is that while Obs:14 agrees with the artist on what constitutes a strong emotion, Obs:6 disagrees, classifying HEE stimuli as low and LEE as strong. In contrast, Obs:13 displays very little tendency to classify the stimuli in a manner resembling the artist with an ROC curve close to that predicted by chance behavior (dashed line).

If the participants were to categorize the stimuli according to the artist’s classification (or in direct opposition, as in the case on the right in Figure 3), it suggests that they detected the signal (i.e., the emotional intensity of the calligraphies). In this case, the signal would be stronger than the noise and the d’ index would be significantly different from 0. Otherwise, participants would be expected to answer randomly in relation to the artist’s classification and the d’ index would not differ significantly from 0.

The rating data were initially fit for each participant individually by a probit regression using cumulative link functions and a maximum likelihood method (McCullagh & Nelder, 1989). The value of d’ is obtained directly as one of the estimated parameters. Standard errors are also estimated which permit testing whether the estimated values differ significantly from zero.

This analysis demonstrated that 16 of the participants had values of d’ significantly different from 0 (see Figure 4), thereby showing that they discriminated the emotional categorization of the stimuli. Among them, four had a negative d’ index. A negative d’ indicates that participants perceived the difference between the two categories of stimuli but identified, in opposition to the artist classification,
LEE as high in emotional intensity and HEE as low in emotional intensity (see Figure 3). For the remaining six participants, d’ did not differ significantly from 0. In other words, these participants did not detect the artist’s classification.

The full dataset was then fit with a mixed-effects model in which we tested simply whether the ratings differed from chance levels of detection. This permitted us to evaluate several potential sources of variability in the data. Specifically, we fit a model that included random effects of Observer (individual differences among participants), Item (individual differences in the difficulty among the 30 stimuli), and Observation (variation within each participant’s rating responses beyond that expected by a multinomial model of rating attribution). The results indicated that the random effects of Observer and Observations were significant (p < .001 in both cases), but that to the random effect of Item was not (p = .06). However, the test employed is conservative and can be too large by as much as a factor of two, which in this case would support significance for an effect of Item as well (Pinheiro & Bates, 2000). The proportion of the total variance accounted for by each source is as follows: Observer (48.2%), Observation Level (49.3%), and Item (2.5%). This suggests that, indeed, that differences in difficulty in rating the different stimuli play only a small part in the overall results.

By treating Observer and Item as random effects, the fixed effect estimates (d’ here) can be extrapolated to a population estimate. The estimated population d’ was 1.076 (SE = 0.223). This indicates a significant resemblance in the classification of calligraphies as a function of emotional intensity level between a non-Chinese speaking population and the artist.

Figure 4. Individual d’ values plotted ordered and with estimated 95% confidence intervals.

Notes: Grey circles indicate non-significant d’ values.
Perimetric Complexity and Amount of Ink of the Stimuli

We investigated whether the physical characteristics of the stimuli could, if at all, be used to categorize the stimuli. Two properties were studied in this regard: (i) the total amount of ink in each calligraphy, and (ii) the perimetric complexity (perimeter squared over “ink” area which is an index that quantifies how convoluted a character is). Perimetric complexity is calculated after having transformed the stimuli into binary images (Atteave & Arnoult, 1956; Pelli, Burns, Farell, & Moore-Page, 2006; Watson, 2012; Watson & Ahumada, 2012). In Figure 5 the two values are plotted against each other with the two types of

![Figure 5. Perimetric complexity.](image)

Notes: The plot depicts in white the HEE calligraphies and in black the LEE ones. The horizontal and vertical lines indicate the optimal separations for either indice (amount of ink and perimetric complexity) alone and the slightly oriented one represents the best line when using both indices at once. The values have been normalized to have equal variance.
images in black for LEE and in white for HEE. Then we asked whether there were linear classifiers that could be drawn through the points that would best discriminate the two sets of points. Once the lines were drawn, we calculated an ideal d’ based on the classification. We found that perimetric complexity (d’ = 2.78) would be a better discrimination cue as compared with total ink (d’ = 1.28) if that is what observers used. If observers were to use both indices together (see the oriented line in Figure 4), however, the d’ value does not improve significantly beyond complexity, indicating that complexity alone would best discriminate the two sets of stimuli.

DISCUSSION

The results of our study demonstrate that emotional properties from Chinese calligraphy can be detected by participants naive to the Chinese language. Naive participants exposed to a series of calligraphic stimuli of different emotional intensity were able to discriminate high emotional from low emotional calligraphy, indicating that the emotional content conferred to the calligraphy by the artist was detected. In other words, what we may call the emotional information expressed in the calligraphy as a result of the calligraphist artistic brushstrokes is reliably extracted. This first result has some implications for the study of emotion perception as well as for the study of the perception of style in Chinese calligraphy.

In terms of emotion perception, participants detected the emotion expressed by the artist independently of the semantic context since they were blind to the Chinese language. A significant number of participants had a positive d’ value. Indeed, the majority were able to separate low emotional intensity calligraphies from high intensity calligraphies. This result is coherent with results from a study that explored emotion expressiveness in music (Gabrielsson & Juslin, 1996). In that study musicians performed several melodies using various instruments so as to communicate specific emotions to listeners. Participants succeeded in decoding the intended expression, leading to the hypothesis that artists and observers share a common code for emotional expression. Similarly, in our study participants grasped the emotional intention of the artist.

Regarding the perception of style in Chinese calligraphy, our results also show that calligraphy can effectively be used as non-representational artistic stimulus in the sense that it can function as an effective means for expression and communication. Emotion detection from calligraphy does not appear to be driven by knowledge. It is possible that low-level properties of the calligraphies suffice to render them emotional. Calligraphy may transmit various emotional intensities through the contours of brushstrokes. As previously mentioned, low-level characteristics may participate in aesthetic judgment, or detection of emotion in stimuli from various categories (Leder et al., 2004). Properties such as contrast, luminance, and color have been shown to play important roles in aesthetic perception (Reber et al., 2004). In this study it was not possible to identify which
properties participants relied on to classify the two categories of stimuli, but there is at least one stimulus cue that was identified as varying between the two sets of stimuli: perimetric complexity.

Complexity has for long been described as one of the features capable of modulating the hedonic value of artistic stimuli (Berlyne, 1970). However, the objective measurement of pattern complexity is rather complicated by its numerous possible definitions. Here we used two measures of complexity: the total amount of ink in each calligraphy and perimetric complexity (Attneave & Arnoult, 1956; Pelli et al., 2006; Watson 2012; Watson & Ahumada, 2012). In general, it appears that LEE stimuli have more ink in them. As for the perimetric complexity, by this measure, nearly all HEE stimuli show more perimetric complexity. We also asked participants, in a post-hoc exploration, to estimate the subjective complexity perceived from each calligraphy on a 5-point scale. Overall, high intensity calligraphies were perceived to be more complex than low intensity calligraphies. Perimetric complexity and, at the subjective level, perceived complexity in calligraphy might have participated in the discrimination between low emotional and high emotional calligraphy.

It is also possible that the detection of complexity was related to the fact that the two sets of stimuli rely on different styles. HEE calligraphy is in a cursive style, whereas LEE calligraphy is in a regular style—this was the artistic choice made by the calligraphist who created the stimuli for the experiment. Since the distinction between HEE and LEE calligraphy does not seem to depend on previously learned knowledge by the participants or their culture, style-related information is reliably perceived in Chinese calligraphy without learning. In our study we have asked participants to see the stimuli as abstract artwork, not as Chinese calligraphy, and to evaluate their perception of emotional intensities or expressiveness of each stimulus. Under these conditions participants may have been able to detect the stylistic properties of the stimulus.

For the philosophy of art, the emotional properties of these two styles of calligraphy may refer to two different notions of emotional expression. We rely on Jenefer Robinson’s (2007) distinction between expression and expressiveness to interpret differences between the two styles of calligraphy in our study. While Chinese calligraphy in the regular and in the cursive styles may both be considered as expressions of emotions, it appears that Chinese calligraphy in the regular style is just minimally expressive and that calligraphy in the cursive style is maximally expressive—in Robinson’s sense of the notion of expressiveness. For Robinson, expressiveness is a matter of degree and is a property of the relation between an artwork and its audience: it “(…) depends on how effectively the artwork reveals to a (suitable) audience what that emotion is like” (p. 36). Since, according to Robinson’s conception of expressiveness, “[a]n artwork that expresses an emotion in an expressive way is one that reveals something of what it is like to be in such an emotional state” (Robinson, 2007, p. 33), we propose that HEE calligraphy expressiveness may be due to its intrinsic evocative
property, whereas LEE calligraphy may be minimally expressive in Robinson’s view because it does not possess this property to a high degree. We propose that the two calligraphic styles used here might correspond to two different modes of emotional evocation.

This is the reason why Robinson’s distinction between expression and expressiveness may shed new light on the distinction between the regular and cursive styles in Chinese calligraphy. It appears that the claim put forward by Cupchik, Vartanian, Crawley and Mikulis (2009, p. 84) on the art of painting “(…) when artists create paintings, (…), they attempt to evoke subjective reactions in the viewer to the stylistic and structural properties of their works” could be applied to the art of calligraphy.

REFERENCES


Direct reprint requests to:

Stephanie Dubal

CNRS USR 3246

Centre Emotion, Pavillon Clérambault

GH Pitié-Salpêtrière

75013 Paris, France

e-mail: stephanie.dubal@upmc.fr.