WHAT MAKES MONA LISA SMILE? INVESTIGATING A HALO MODEL OF CONFIGURAL FACE EFFECTS

Armando M. Oliveira[1], Ana Duarte Silva[1], Nuno Sá Teixeira[1], Ricardo G. Viegas[1]
[1]Institute of Cognitive Psychology_University of Coimbra, Portugal
l.dinis@fpce.uc.pt

Abstract

Da Vinci's Mona Lisa is often celebrated for her enigmatic expression, somewhere in between happy and sad. Changes in Mona Lisa's mouth (by superimposing noise on it) have been shown to affect separate ratings of the eyes regarding their happy or sad character, despite these being physically unaltered. Like many other configural face effects, this one has been taken at face value as illustrating a mouth-eyes interaction. The present study investigates an alternative explanation based on the weighted averaging halo model of IIT (Anderson, 1981), which assumes no interaction among components.

The issue of context effects has always been present in the study of faces. Typically, context information makes reference to context outside the face, acknowledging the fact that faces seldom stand in isolation but are commonly met instead in rich interactive settings. The ubiquity of outer contexts also draws attention to the fact that, however important as a communicative device, the face is one of manifold channels of communicative signs (e.g., postures, gestures, prosody, proxemics, verbal utterances, etc.). By far, the dominant question posed by research has concerned the relative importance of contextual information (regarding that of face) for the judgment of emotions (Goodnough & Tinker, 1931; Frijda, 1969; Ekman, Friesen & Ellsworth, 1982; Nakamura, Buck, Kennedy, 1990; for reviews see Wallbott, 1988; Fernandez-Dols & Carroll, 1997). This is also true of the multichannel perspective. Establishing the relative importance of the face channel vis-à-vis other channels, such as voice quality or intonation, has been a standard topic of research along this line (Meharabian & Ferris, 1967; Ekman et al., 1980; Hess, Kappas, & Sherer, 1988). One problem faced by this stream of inquiry, which mostly rests on correlation and regression methods, is lack of valid means for measuring importance (see Anderson, 1982, pp. 262-277; 2001, pp. 275-279, 557-559). Virtually all conclusions drawn over the relative influence of context are thus unwarranted (Anderson, 1989, 165-167). The forthcoming section will illustrate instances of this problem and the solution provided by FM, owing to the weight-value distinction afforded by the averaging model.

However, a more fundamental problem than the one of measurement is whether context and facial expression can be meaningfully addressed as separate informers. Attempts at measuring their respective importance implicitly assume that they do, but this assumption has been questioned on the ground that face and context actually change each other's meaning when in combination (Fernandez-Dols, & Carrol, 1997). The main argument for that rests on phenomena of «vulnerability to reinterpretation» (Fernandez-Dols & Carrol, 1997) whereby changes in context (face) lead to a reconsideration of the meaning attributed to the face (context). Such effects (of which the often cited Kuleshov effect in film editing is an example) actually do little more than comfort the phenomenal impression that contextual information changes the meaning of a face, with no bearing upon the mechanisms whereby context exerts its effect. The interactive interpretation of the face-context relation thus remains no less an assumption than the non interactive, independent account.

Contrary to the idea that this is a strictly empirical matter, distinguishing between the two interpretations cannot be done without adequate model analytic capabilities. This is a documented fact within the FM program, which very early on set forth an alternative to the meaning-change interpretation of context effects (Anderson, 1981, pp. 161-169). This alternative tack involves a two-step integration: all informers (e.g., face and context) are first combined into an overall impression *I*; *I* is then integrated with the particular informer under evaluation (e.g., the face). At the end, the face has undergone a change in phenomenal value, but no intrinsic interaction has occurred between face and context. In a further specification, the integration between *I* and the judged informer was admitted to obey a weighted averaging rule:

$$s' = ws + (1 - w)\,, \tag{1}$$

with *s'* the rating of the in-context informer, *s* its free-context value, and *w* its relative weight in the integration. This formulation corresponds to the 'averaging halo model' of IIT, so called because it rests on the influence of an overall impression over the evaluation of a particular component (Anderson, 1981, pp. 235-244; 1996, 112-115). One noteworthy point is that, as all IIT/FM models, the halo model only assumes the meaning invariance of informers relative to a task-dependent goal and to a judgment dimension, not as a general property of the stimulus. So, «vulnerability to reinterpretation» in the sense argued by Fernandez-Dols ultimately agrees with the FM view that no fixed psychological value preexists in the stimulus. In the meanwhile, it entails no consequence as to how context and the face integrate to produce a judgment along a given response dimension, which is the issue of the meaning-change versus the halo interpretation debate.

Components in a face may be regarded as the surrounding context for a particular component being judged on some dimension. The present study is concerned with such context effects inside, and not outside, the face. It started off from work by Kontsevich and Tyler (2004) over the Mona Lisa's smile. Using the *sfumato* technique, Da Vinci managed to produce an elusive smile which is best seen when not looked at directly (see Livingston, 2000) and tinges the face with an enigmatic expression, somewhere in between happy and sad (Figure 1, leftmost image). By overlaying noise on Mona Lisa's mouth (which produced a change in the perceived smile towards either the happy or the sad pole) Kontsevich and Tyler documented consistent effects upon separate ratings of the eyes on a happy-sad dimension, despite these being physically unchanged (see middle and rightward images in Figure 1, panel A).

A B

Figure 1. *A.* Mona Lisa's smile configural effect: component eyes (unchanged) are perceived more on the happy pole (middle face) or on the sad pole (rightward face) as a function of mouth changes. The leftward face is a gray scale facsimile of the original Mona Lisa's face. *B.* Reproduction of the same configural effect in a synthetic face.

Like many other configural face effects, this one was taken as illustrating a mouth-eyes interaction. However, an alternative explanation exists through the averaging halo model, which assumes no interaction among components. This was investigated in two experiments with synthetic faces (see Figure 1, panel B). In one of them, eyes/brows-related and mouth-related AUs varying along a happy-sad continuum (five levels each) were factorially combined to produce the stimuli-faces. Besides ratings of whole facial expressions on a bipolar happy-sad graphic scale, mouth and eyes components were also separately rated. Thus, when eyes were being evaluated, mouth acted as a context for the evaluation, and vice-versa. The other experiment was similar, except that AUs varied along an anger-happy continuum. Three conditions were additionally created within each experiment, concerning the way faces were presented: up-right, inverted, and with misaligned top and bottom halves. The two later manipulations were expected to increasingly reduce the magnitude of contextual effects, for reasons given below.

Two sorts of predictions from the averaging halo model were tested in the data. The first one concerns the linear relation between the ratings for the eyes or mouth component and the ratings for the overall expression (I), as expressed in the equations

$$eyes' = w_e eyes + (1 - w_e)I;\ mouth' = w_m mouth + (1 - w_m)I, \tag{2}$$

with *eyes'* and *mouth'* the in-context judgments, and w_e and w_m the weights of the free-context *eyes* and *mouth*, respectively. One facet of this linear relation is that overall shape of the patterns for the whole expressions should be reflected in the ratings of components. In case the impression I itself arises from a linear integration rule, parallelism should then result in the components judgments (Anderson, 2001, 236; Anderson, 1996, 114). However, in case I presents systematic deviations from parallelism, those same deviation trends should be mirrored in the contextual ratings.

Results presented in Figure 2 (contrasting judgments of the overall expression, I, and of the eyes/brows across presentation conditions) overall agree with this prediction, as revealed by vertical comparisons within each column.

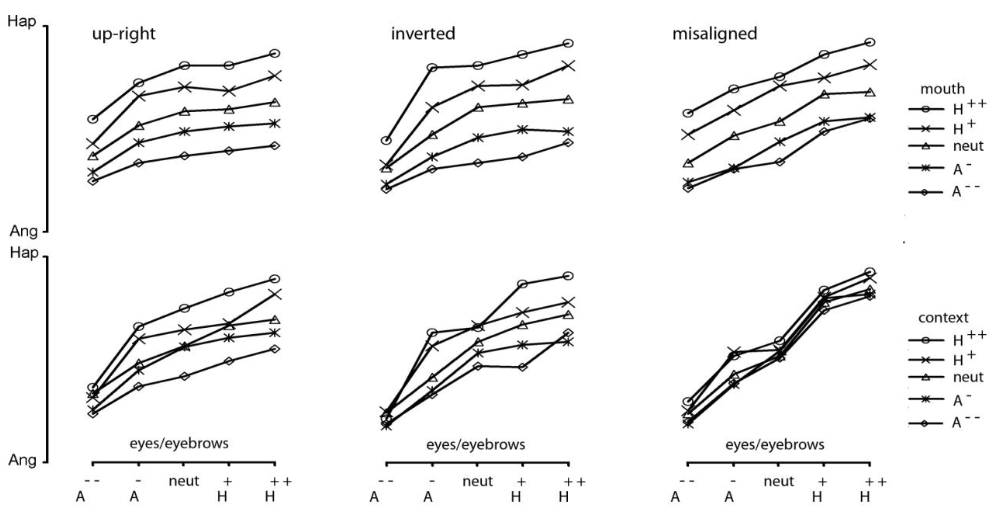

Figure 2. Factorial plots for judgments of the whole expression (top row) and of the eyes/brows component (bottom row) in the Anger-Happy experiment. Mean ratings are on the ordinate, corresponding to a bipolar anger-happy scale (most anger at the bottom, most happiness at the top). Columns correspond to the three presentation conditions.

The extremity weighting effect observed for the leftmost plots in the top row (downward convergence of lines, associated with the 'negative' anger levels) is well replicated in the bottom row. Also, except for specific points, the pattern of inner spacing between lines shows evidence of rough proportionality between judgments of overall expressions and of eyes-in-context. The misaligned condition has the particular property that it exhibits parallelism (sign of a linear integration) for the general expression. According to predictions, parallelism should also emerge in the component judgments, which it does (despite a dramatic reduction of the context effect in this condition, it still reached statistical significance).

The second prediction concerns the effects of presentation mode. The inner relatedness of facial components is widely believed to decrease with face inversion (Yin, 1969) and still more with misalignment of top and bottom face halves (Young, Hellawell, & Hay, 1987). This can be expected to impact on the relative weight of the overall impression (1-*w*), while the context variable affects *I* itself (see Anderson, 1981, 244).

From the $(1\text{-}w) \times I$ term in the halo equation it can thus be predicted that plotting context against presentation mode for each level of the rated component will result in graphic linear fans, expressing this multiplicative relation (Anderson, 1981). Plots in Figure 3 strongly support this prediction. Not only do context effects decrease from the upright to the inverted to the misaligned condition (in the expected order, thus), overall they comply with the line fanning typical of multiplying models. Being the less affected by context, the misaligned condition (dashed line) can be seen to operate as a baseline for the fanning. The inverted condition always displays a lesser slope than the up-right condition, concurrent with a decreased relative weight (1-*w*) of the overall impression. These results agree with those obtained by Takahashi (1971; see p. 172) in the domain of personality impression formation. Even if he interpreted them as being against the averaging halo model, they can actually be predicted from it, as pointed out by Anderson (1981, 244). One entailed consequence for holistic processing is that striking examples of holistic effects can actually dispense with configural interaction (essential configurality) and be accounted for in terms of algebraic, non interactive configurality.

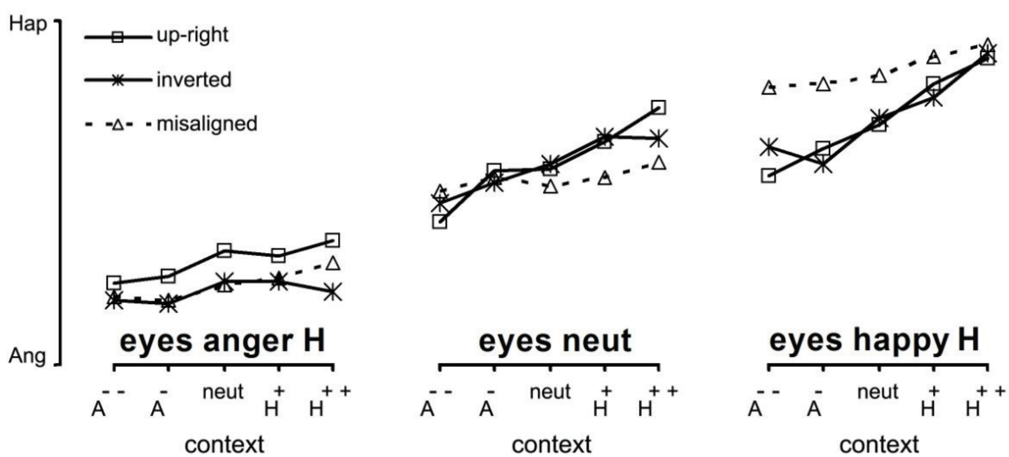

Figure 3. Factorial plots for context (abscissa) × presentation mode (curve parameter). Each graph corresponds to a particular level of the eyes component being targeted for judgment (outcomes presented for only three of the five levels (mean ratings on the ordinate). The context factor is represented by the five levels of the mouth factor, ranging from intense Anger (A++) to intense Happiness (H++).

References

Anderson, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.

Anderson, N. H. (1982). *Methods of information integration theory*. New York: Academic Press.

Anderson, N.H. (1989). Information integration approach to emotions and their measurement. In R. Plutchik and H. Kellerman (Eds.) *Emotion: Theory, research, and experience. Volume 4: The measurement of emotions* (pp. 133-186). San Diego, CA: Academic press.

Anderson, N. H. (1996). *A functional theory of cognition*. Hillsdale, NJ: Erlbaum.

Anderson, N. H. (2001). *Empirical Direction in Design and Analysis*. Hillsdale, NJ: Erlbaum.

Ekman, P., Friesen, W. V., O'Sullivan, M., & Scherer, K. (1980). *Relative Importance of Face- Body- And Speech In Judgment_of personality and affect.* Journal of Personality and Social Psychology. 1980, 38, 270-277.

Ekman, P., Friesen, W., & Ellsworth, P. (1982). What emotion categories or dimensions can observers judge from facial behavior? In P. Ekman (Ed.), *Emotion in the human face* (pp. 39-55). New York: Cambridge University Press.

Fernandez-Dols, J. M., & Carroll, J. M. (1997). What does a facial expression mean? In J. A. Russell & J. M. Fernandez-Dols (Eds.), *The psychology of facial expression* (pp. 275-294). Cambridge, UK: Cambridge University Press.

Frijda, N. H. (1969). Recognition of emotion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 4, pp. 167–224). New York: Academic Press.

Goodenough, F. L., & Tinker, M. A. (1931). The relative potency of facial expression and verbal description of stimulus in the judgment of emotion. *Comparative Psychology*, *12*, 365-370.

Hess, U., Kappas, A., & Scherer, K. R. (1988). Multichannel communication of emotion: synthetic signal production. In K. R. Scherer (Ed.), *Facets of emotion: Recent research* (pp. 161-182). Hillsdale: NJ: Erlbaum.

Mehrabian, A., & Ferris, S. R. (1967). Inference of Attitudes from Nonverbal Communication in Two Channels. *Journal of Consulting Psychology, 31(3)*, 248–252.

Nakamura, M., Buck, R., & Kenny, D. A. (1990). Relative contributions of expressive behavior and contextual information to the judgment of the emotional state of another. *Journal of Personality and Social Psychology, 59(5)*, 1032-1039.

Takahashi (1971). Effect of the context upon personality-impression formation. *The Japanese Journal of Psychology, 41*(6), 307-313.

Wallbott, H. G. (1988). In and out of context: influences of facial expression and context information on emotion attributions. *British Journal of Social Psychology, 27*, 357–369.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*, 141-145.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, *16*, 747-59.

Acknowledgments

This work was supported by Grant PTDC/PSI/73406/2006, from the Portuguese Foundation For Science and Technology