

CHAPTER 14: THE MEANING OF THE LEFT CHEEK

Summary

The meaning of portraits showing either the left or the right cheek has been studied by means of Semantic Differential technique.

The particular cheek presented to the subject produces little difference in meaning. There are however significant differences in meaning between those portraits originally painted showing the left cheek as opposed to the right cheek.

14:1 Introduction

Portraits tend to show the left cheek more often than the right (see Chapter 13). The difference is more pronounced in portraits of women than of men. For self-portraits however it is the right cheek that is more common. On the basis of the latter finding, and in conjunction with a detailed analysis of Rembrandt's work, Humphrey and McManus (1973) suggested that the meaning of the two cheeks was different, and that when the artist painted an individual showing the right cheek then he was portraying that person as being closer to the artist than if the left cheek was painted; this hypothesis was particularly supported by the relation between the cheek shown in Rembrandt's portraits and the kinship status of the subjects. The hypothesis has since received further support from a detailed analysis of paintings of the Crucifixion, Annunciation and Madonna and Child in Mediaeval and Italian Renaissance art, (Chapter 13), and of the same picture types in Flemish, Dutch and German paintings of the Renaissance and Reformation (Wolfe, 1978).

The question thus arises whether judgements of an ordinary subject, when viewing a portrait, are influenced by the cheek shown. The present study has investigated this question by showing portraits and their mirror-images to a number of subjects and asking them to make semantic differential judgements of the stimuli.

14:2 Experiment 1Method

Twenty subjects were shown a series of 16 portraits of single individuals, and asked to rate the person in each portrait on 20 semantic differential scales. The 20 scales for each portrait were each placed on an individual sheet in an answer booklet. Each successive sheet in the booklet was the mirror-image of the previous one; thus if sheet 5 had Good-Bad at the top, sheet 6 had Bad-Good, and sheet 7 was the same as sheet 5. 10 of the subjects started with one version of the first sheet, and the other ten with the other version. Each scale had seven divisions on it, and subjects were instructed to use the scales as in the instructions on pages 82-84 of Osgood et al (1957).

Of the 16 portraits seen by each subject exactly half showed the left cheek and half the right cheek. Each portrait was shown to half the subjects showing its left cheek and to the other half of the subjects showing its right cheek (by reversing the slides in the projector). There were thus two series of portraits, the order in one being randomised, and the other being the mirror-image of the first series. Half of the portraits were of males, and half were of females. Each portrait showed only a single person. The stimuli were photographed from post-cards and back-projected on to a ground-glass screen in a dimly

illuminated room. The stimuli were of the order of 40 x 40 cms, and were viewed from a distance of about 1.5 metres. The rate of presentation of the stimuli was controlled by the subjects, by means of a push-button. No time limits were set; most subjects completed the experiment within 30-45 minutes. All subjects were tested individually. The order of the stimuli, left and right cheek, male and female, was randomised, with the constraint that one of each of the four possible stimulus types occurred in the first four stimuli.

Of the 20 subjects, 8 were left-handed and 12 were right-handed, and 14 were male and 6 were female. Subjects were mostly undergraduates of the University of Cambridge, and none were reading art, history of art or architecture. None were aware of published work on the left cheek phenomenon.

Having carried out the first part of the experiment the subjects were then shown a further four slides, these being repeats of the first four slides shown in the experiment. The subjects were informed that as the first four sets of judgments in an experiment were often unreliable, the purpose of the repetition was merely to allow a more accurate response to the first four stimuli. Six of the subjects did indeed see the first four slides as they had originally been shown, thus allowing a measure of test-retest reliability. The remaining fourteen subjects

saw the first 4 stimuli in the mirror-image of the form in which they had first seen it. Only two subjects noted the reversal, and they noticed it only for a single stimulus each. This part of the experiment thus allowed a within subject analysis.

After the experiment the subjects were asked what they thought the experiment was intended to discover. No-one mentioned any aspects of right or left. The subjects were then informed of the true purpose of the experiment.

Results

The judgements on the 20 scales (shown in Table 14:1) were factor-analysed in order to reduce the number of variables to be considered. This was carried out individually for each subject (since there was concern that there might be large individual differences in semantic space), and also for the composite data of all subjects. No important differences were found between the two methods, although there were slight inter-subject variations; the population factor-analysis was used for the results in this paper. The judgements on each of the 20 scales were inter-correlated and the resultant 20 x 20 matrix factor analysed by means of the SPSS programs (Nie et al, 1975), using the PA2 option, and then a Varimax rotation. The first three factors had eigen-values of 5.9, 3.8 and 2.1, all highly significant. The next five eigen-values were 0.96, 0.90, 0.83, 0.70 and

0.64, all of which are clearly non-significant, and thus only the first three factors, accounting for 59.6% of the total variance, were rotated. Table 14:1 shows the 20 scales, their communalities, the loading on each of the varimax rotated factors, and the per cent of the common variance attributable to each factor.

To obtain three factors in an experiment such as this one is fairly typical (Osgood et al, 1957). Naming of the factors presented little difficulty. Factor 2 clearly corresponds to Osgood's 'Evaluation', and Factor 1 to Osgood's composite Potency and Activity scale, or 'Dynamism'. Factor 3 does not relate to Osgood's analysis and is clearly a function of the particular scale items chosen in the present experiment, items such as Spiritual and Sacred being included to test the possibility (derived from studies of Italian Renaissance painting), that the left cheek might be more spiritual; Factor 3 has been labelled as 'Spirituality'.

From the factor analysis the judgements of each picture by each subject were recalculated as scores on each of the three factors (these scores each have an overall mean of zero and variance of unity, although this does not apply to the results of individual subjects).

Because of the design of the experiment, by which each subject saw four of the pictures on two separate occasions,

it is possible to produce an error estimate for each subject. However a complete two-way analysis of variance (i.e. subject and picture plus subject picture) is not possible. Considering just main effects (i.e. ANOVA model is subject plus picture) there are highly significant differences in Evaluation, Dynamism, and Spirituality for different pictures ($F(15,365) = 12.50, p < 0.001$; $F(15,365) = 2.42, p < 0.001$; $F(15,365) = 14.41, p < 0.001$ respectively) and also for different subjects ($F(19,365) = 2.415, p = 0.0009$; $F(19,365) = 2.375, p = 0.0011$; $F(19,365) = 1.613, p = 0.0505$ respectively). Since overall differences in means between subjects are not of real interest, these differences have been removed by standardising each subject's scores so that they have a mean of zero, and variance of unity. It is now possible to examine subject-picture interactions by using an ANOVA model of the form (picture + picture \times subject). Cheek shown to the subject may also be analysed as a main effect within this model; this produces a model of the form 'picture + picture subject + cheek shown', the three levels having 15, 304 and 1 df respectively, there being a residual with 79 df's.

Picture main effects are still all significant with probability very much less than 0.001. Picture subject interactions for each of Evaluation, Dynamism and Spirituality are all significant ($F(304,79) = 2.408, p < 0.0001$; $F(304,79) = 2.380, p < 0.001$; $F(304,79) = 2.379, p < 0.001$ respectively).

Thus there are differences between the ways in which particular subjects view particular pictures.

The cheek shown does not seem to affect the perception of the picture ($F(1,79) = 0.211$, $p = 0.647$; $F(1,79) = 0.086$, $p = 0.770$; $F(1,79) = 0.289$, $p = 0.592$ respectively).

From experiment 1 we must therefore conclude that it is unlikely that differences exist in the evaluation of pictures according to the cheek shown to the subject.

The hypothesis of McManus and Humphrey (1973), regarding the meaning of the left and right cheek, seems to receive no support from the experiment so far. However the original hypothesis primarily concerned the perception of pictures by artists. It is thus possible that if we carry out an experiment with the original artists, then significant results would be obtained. In the absence of this, an a posteriori analysis of the data of experiment 1 was carried out to discover whether the judgement of a portrait related, not to the cheek as perceived by the subject, but to the cheek as originally painted by the artist. Clearly there is no possibility of a within subjects analysis of such an hypothesis since it is not possible to obtain a second, mirror-imaged version of the painting, as painted by the artist. One may therefore carry out only a between subjects analysis.

Table 14.2 shows, for the total results of experiment 1 the difference in mean score on each of the three factorial dimensions between those pictures showing the left cheek in the original ($n = 10$), and those showing the right cheek ($n = 5$); (one picture could not be classified in this manner as the original could not be found). For evaluation and dynamism the p levels in a one-way analysis of variance (two-tailed) show a tendency, particularly in the case of dynamism, for the two portrait types to be different. The scale in Table 14.2 is SD (z) units, and hence the left cheek pictures appear to be half of a z unit less dynamic or evaluatively positive. Clearly very little can be made from this result since the analysis is a posteriori, and the sample sizes are small. Hence a second experiment, with more pictures and less subjects was designed to study the question further.

14:3 Experiment 2

In this experiment subjects were shown a series of 106 portraits, and asked to rate the portraits on a semantic differential consisting of a subset of nine of the scales used in experiment 1 (marked with an asterisk in Table 1); these scales were selected for their high loadings on the three main factors. Subjects saw the total of 106 portraits in four separate experimental sessions, the first three sessions consisting of 26 portraits, and the fourth of 28 portraits. Sixty-six of the portraits,

1702

originally showed the left cheek, and 40 the right cheek (approximately the population proportions). In the series of 106 portraits half were presented to the subject showing the left cheek, and half the right cheek, the orientation being determined independently of the original cheek painted by the artist.

Four subjects participated in the experiment, and each took from 25 to 60 minutes on a session. The experimental situation was basically similar to that of experiment 1, except that stimuli were front projected, and were slightly larger. Two of the subjects saw one particular randomised sequence of the 106 portraits, and the remaining two subjects saw the mirror image of this stimulus sequence.

In view of the small number of subjects the data for each subject were factor-analysed individually, using the same method as described in experiment 1. Factor-scores were then calculated for the judgements of each portrait on each of the three dimensions. In each subject the three dimensions of Evaluation, Dynamism and Spirituality were readily identified, although there were minor differences in emphasis and correlation pattern between subjects. For each subject the distribution of factor-scores on each dimension had a mean of zero and a variance of unity.

As with experiment 1, no differences could be found in experiment 2 between pictures according to the cheek

actually shown to the subject, although this was not surprising in view of the between subjects design, and the negative result of experiment 1. Detailed results will not be reported here.

Table 14:2 shows the results of both experiments according to the cheek originally painted by the artist. The a posteriori results of experiment 1 have been taken as indicating the expected direction of any effect in experiment 2, and thus the table shows, separately for each subject, the one-tailed probability of a difference between right and left cheek. Results for the four subjects have been combined using the method of Kendall (1951). For Spirituality and Dynamism there is no evidence of a replication of the effects of experiment 1, ($\chi^2_8 = 2.96$, and $\chi^2_8 = 11.63$ respectively, both NS). For Evaluation, three of the subjects show an effect in the same direction as experiment 1, and for one of these subjects the result is highly significant ($p = 0.0007$) and for another the result is almost significant ($p = 0.06$). Combining the four subjects by the method of Kendall (1951) produces a χ^2 value of 25.97, which with 8 df is significant ($p < 0.001$). As in experiment 1 this result cannot be put down to the effects of repeated comparisons, as if one combines the three overall values for the four subjects, one still obtains a χ^2 value of 40.56, which with 24 df is still significant with $p = 0.0186$.

Lest it be thought that the statistical method of analysing the subjects separately and then combining them is not valid it may be pointed out that if the original scale judgements of each subject are combined and then these composite results factor-analysed and the resultant factor scores then analysed in terms of the original cheek painted by the artist, then the only significant result is for evaluation, with a one-tailed significance of $p = 0.0215$.

Experiment 2 thus provides a replication of the most surprising finding of experiment 1, that the evaluation of a portrait depends not upon the cheek as shown to the subject, but upon the cheek as painted by the artist. The magnitude of the effect is slightly smaller in experiment 2 than in experiment 1, the right cheek being 0.215 SD units more evaluatively positive, as compared with 0.580 SD units in experiment 1.

14.4: Discussion

The results of this experiment are initially surprising; a subject's evaluation of a portrait is not dependent upon the cheek as perceived by him, but instead relates to the portrait as conceived originally by the artist. From judgements made independently of the asymmetry of a portrait we may gain knowledge as to its original asymmetry, a superficially counter-intuitive result.

Two extremes of explanation arise, although intermediate positions are possible. Despite the lack of any significant results in the first part of experiment 1, it is nevertheless still possible that there is a gradient of sensitivity to the meaning of cheeks, ordinary subjects hardly being sensitive at all, and artists being very sensitive.

An alternative explanation is that indeed the ordinary subject has no sensitivity to the cheek shown, and the artist himself is only sensitive to it in so far as it is related to an iconographic tradition. The ordinary subject would thus have no access to such restricted information, and hence could make no assessment at all on the basis of the perceived cheek.

However in neither case is it unreasonable to postulate that the ordinary subject's judgements will be related to the original cheek, as painted. Since it seems to be typical (or perhaps even necessary) for stimuli to be evaluated in three semantic dimensions only (Osgood et al, 1957), then the artist's sensitivity to the cheek would be linked on to one of the three semantic dimensions. However, there are many other cues as to these dimensions apart from the cheek, and since the subject would have access to this alternative information about the dimensions the subjects judgements would be expected to relate to the cheek as originally painted. If correct this hypothesis would predict that assessments by artists of other artists paintings might depend for reasons the

rather than upon the conceived cheek.

Whether or not the ordinary subject can make any judgement at all of the perceived cheek is of importance in determining the origin of the left cheek phenomenon. If the phenomenon has biological origins, perhaps in the asymmetry of the human brain, we would expect all subjects to show the phenomenon to a greater or lesser extent, although artists would be expected to show the phenomenon far more strongly. If however the phenomenon is primarily iconographical (and hence, in some sense, arbitrary), then ordinary subjects should show no awareness of the phenomenon at all. The present experiment suggests that the iconographical interpretation may well be correct, although a larger replication of the main result of experiment 1 is necessary before this question can be answered definitively.

In summary the meaning of a picture is clearly related to the cheek shown in the original version, but is unlikely to be related to the cheek as perceived.

Table 14:1

Shows the 20 semantic scales of experiment 1, their communalities and loadings on the three Varimax factors. An asterisk next to a scale indicates that it was used in experiment 2.

NOTE: Decimal points have been omitted from all loadings and communalities. Loadings which may be regarded as significant (i.e. ≥ 0.30 or ≤ -0.30) have been underlined.

For each pair of words in a scale the first one is the one referred to in the loadings.

Scale	Dynamism Factor 1	Evaluation Factor 2	Spirituality Factor 3	Communality
*Good-Bad	08	<u>72</u>	33	65
Noble-Ignoble	<u>42</u>	24	<u>39</u>	39
*Wise-Foolish	<u>70</u>	18	27	61
*Kind-Cruel	-18	<u>80</u>	22	74
Beautiful-Ugly	-02	<u>68</u>	10	48
Successful-Unsuccessful	<u>78</u>	<u>-06</u>	-05	62
Important-Unimportant	<u>74</u>	-16	12	60
*Strong-Weak	<u>75</u>	-12	-22	63
*Remote-Intimate	26	-62	28	53
Masculine-Feminine	<u>39</u>	<u>-38</u>	-06	30
Hard-Soft	<u>54</u>	<u>-68</u>	-23	82
*Excitable-Calm	<u>-36</u>	-13	<u>-41</u>	31
Active-Passive	<u>64</u>	-12	<u>-38</u>	58
*Powerful-Powerless	<u>85</u>	24	-15	81
Tragic-Comic	-09	-01	<u>39</u>	16
*Spiritual-Physical	-04	23	<u>70</u>	54
Personal-Impersonal	-24	<u>74</u>	-08	60
*Sacred-Profane	-01	<u>44</u>	<u>55</u>	50
Eerie-This worldly	-05	-11	<u>50</u>	26
Like me-Unlike me	20	<u>53</u>	02	33
% Common Variance	52.0%	32.1%	15.9%	

Table 14:2 Results of Experiment 2, with results of Experiment 1, for comparison.

<u>Difference between Cheeks ($\bar{L}-\bar{R}$)</u>			
	<u>Evaluation</u>	<u>Dynamism</u>	<u>Spirituality</u>
<u>Experiment 1</u>			
$(\bar{L}-\bar{R})$	-0.5801	-0.4991	-0.2749
Two-tailed p (F-test)	p = 0.1537	p = 0.0851	p = 0.2686

<u>Experiment 2</u>			
<u>Subject 1 (WK)</u> ($\bar{L}-\bar{R}$)	-.2963	-.0142	-.2990
One-tailed p (F-test)	p = 0.0600	p = 0.471	p = .0545
<u>Subject 2 (AH)</u> ($\bar{L}-\bar{R}$)	+.0106	+.2099	+.0790
One-tailed p (F-test)	p = 0.5225	p = 0.858	p = 0.667
<u>Subject 3 (JMF)</u> ($\bar{L}-\bar{R}$)	-.5369	+.0456	-.2258
One-tailed p (F-test)	p = 0.0007	p = 0.598	p = 0.1095
<u>Subject 4 (RFD)</u> ($\bar{L}-\bar{R}$)	-.0385	+0.2725	+.1198
One-tailed p (F-test)	0.4179	p = 0.9395	p = 0.749
<u>Combined results of Experiment 2</u>			

One-tailed tests			
X^2 (8 d.f.)	25.97	2.96	11.63
P	(p = 0.0011)	(p = 0.936)	(p = 0.167)