# THE EVALUATIVE JUDGEMENT CONTINUA OF EXTREME AND NON-EXTREME PATIENT RATERS

A thesis presented in the Department of Psychology and Sociology
University of Canterbury

In partial fulfillment of the requirements for the Degree of Master of Arts

by

PERCY JOHN BOYES

January, 1966

ACKNOWLEDGEMENTS

The author wishes to acknowledge the guidance and constructive criticism of his supervisor Dr. A.Z. Arthur. Thanks are also due to the Medical, Psychological (particularly Mr. H.E. Bernhardt), and Nursing Staff of Sunnyside Hospital; and also to Dr. R.A.M. Gregson for advice on statistical treatments.

Without the participation of the patients of Sunnyside Hospital this study would not have been possible; they are thanked for their co-operation.

TABLE OF CONTENTS

# CHAPTER 1

## INTRODUCTION AND AIM OF THE STUDY.

There has been a tremendous amount of psychological literature dealing with response sets, biases, and styles in the last few years (Jackson & Messick 1958; McGee 1962; Rorer 1965; O'Donovan 1965). That such a phenomenon exists seems pretty well established although Rorer (1965) has cast some doubts on the methodological adequacy of many of the research designs used. A large part of the research reported has been concerned with the response biases or test taking attitudes in personality, interest and attitude inventories, where the subject shows a more or less consistent tendency to make a particular kind of response to a test.

Biases have variously been described by such terms as "social desirability" (Edwards 1957b, 1959), "defensiveness" (Smith 1959), "self-disclosure" (Jourard 1959), "yeasaying" and "naysaying" (Couch & Keniston 1960), "deviation" (Berg 1955, 1961), and "extreme position response bias" (Arthur 1966).

Arthur (1965a) in his review found four different approaches used in explaining response bias. They include Cronbach's (1950) emphasis on the form and content of the test; the statistical approach of Cronbach and

Gleser (1953) which was developed by Berg (1961) and Barnes (1955) who have suggested that the individuals who are deviant from the norm on any one response measure tend to be deviant on a variety of statistically normative measures; the contributions of personality exemplified by Edwards (1957b); and the importance of human behaviour as a cognitive activity put forward by Arthur (1965a). This last mentioned takes up a not yet fully explored point of view which was first suggested by Osgood et al., (1957), who pointed out that some subjects are extreme responders on the Semantic Differential, that is, they tend to check the extreme positions 1 and 7 more frequently than they check the other scale positions. They considered that intelligence, age, emotionality, response conflict, anxiety, mental disorder (pp 226-236), and intensity of mediating reactions (pp 155-159) were some of the variables which contributed to the response style.

While there is no evidence that intelligence affects response style (Ware 1958; Neuringer 1963), it has been shown that age is important, children and old people tending to make more extreme responses (Arthur 1965b); (Donahoe 1962). It is not clear at present to what

extent the other factors suggested by Osgood et al. (1957) are independent variables.

Arthur (1965a) in putting forward explanatory hypotheses, found that the form of the semantic differential tends to encourage extreme responses in psychiatric patients. This finding reduces the significance of the extreme response bias in a judgement situation and suggests that some portion of the bias is the result of the specific test form rather than of judgement. This conclusion is in line with Cronbach's (1950) observation that response sets become more influential as the test items become more difficult or ambiguous. In addition Arthur (1965a) found that the failure of discrimination hypothesis put forward by Kelley et al. (1955) was not tenable when applied to the semantic differential in a psychiatric setting.

The present study is an extension of Arthur's (1965a) work and its aim is to throw light on the nature of the extreme response bias. It attempts to show first, that psychiatric patients can be differentiated into two groups in terms of extreme responses in an absolute judgement situation (semantic differential), and secondly, that the groups so differentiated will also differ in

their judgements as shown by a scaling of their responses in a paired comparisons judgement situation. It was expected that if extreme responders were characterized by some different processes of judgement from non-extreme responders, then in a paired comparisons situation where extremeness is not allowed by the nature of the task, they might allocate different values to the concepts to be judged or use different scale widths, and these would be measured by scaling their responses.

## CHAPTER 2

PROBLEM.

The first problem is to find whether patients differentiated in terms of extreme responses fall into a natural group; that is whether their intermediate responses are non-overlapping with non-extreme while they do not differ in the use of neutral positions. There is well documented evidence that psychiatric patients tend to be extreme raters (Lewis & Taylor 1955; Berg & Collier 1953; Zax et al., 1964; Barnes 1955; Borgatta & Glass 1961). It has also been shown that among patients, psychotics tend to be more extreme than neurotics (Arthur 1966; Parsonson 1965), and neurotics more than normals (Wertheimer & McKinner 1952), thus it is expected that in an unselected group of psychiatric patients there would be a wide range of responses. The standard semantic differential was chosen as a measure of extreme response because both Arthur (1966) and Parsonson (1965) have shown that this instrument can differentiate certain psychiatric patients in terms of response bias. In addition it may be of clinical value to know the nature of response bias in the semantic differential which measures meaning. Using an unselected group of psychiatric patients, but excluding mental defective and

older patients over 50 years on the grounds that they might not understand sufficiently the nature of the task, it was hypothesised that two groups would be identified by response bias, these groups to be labelled "Extreme" and "Non-extreme" respectively.

The second problem was to measure the evaluative judgement continua of the two groups differentiated by response bias on the semantic differential. It was thought desirable to keep the task as close to the semantic differential as possible and yet at the same time eliminate the extreme position response 'per se'. To this end the method of paired comparisons was chosen, enabling a patient to make judgements in terms of each pole of the evaluative dimension, good and bad, at separate times. Thus, not only would patients judge which of two concepts was "better", but at a different point in time they would judge which of those two concepts was "worse". It was hypothesised that the Extreme and Non-extreme groups would differ in their underlying scaling behaviour on an evaluative dimension where extreme judgements were not possible.

Osgood et al. (1957) have assumed that the adjective pairs they used in their semantic differential are

bipolar or symmetric around a neutral point. On the basis of this assumption it was hypothesised that the scaling of the "worse" judgements from a given group when reversed would not differ significantly from the "better" judgements for the same group. This condition would therefore serve as a replication.

## CHAPTER 3

## METHOD.

### THE SEMANTIC DIFFERENTIAL

For purposes of differentiating the extreme raters from non-extreme raters three major dimensions of "Evaluation", "Potency" and "Activity", found by Osgood et al. (1957) in their factor analysis were represented in the semantic differential scales used (see Table 1). The seven standard concepts used were adapted from those employed in the Thesaurus sampling of the semantic differential by Osgood et al. (1957) and also by Arthur (1966). The concepts were always judged in the order presented below:

Flower

My Mother

Boulder

Sin

Symphony

New Zealand

Myself

Form II of the semantic differential was used allowing one concept to each page, all the scales being set out below the concept (see Table 1). The order and direction from left to right of the ten scales was

balanced out for each dimension and was maintained throughout the series.

TABLE 1

Example of a page of the Semantic Differential

FLOWER

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| kind | ____: | ____: | ____: | ____: | ____: | ____: | ____: cruel |
| passive | ____: | ____: | ____: | ____: | ____: | ____: | ____: active |
| good | ____: | ____: | ____: | ____: | ____: | ____: | ____: bad |
| weak | ____: | ____: | ____: | ____: | ____: | ____: | ____: strong |
| true | ____: | ____: | ____: | ____: | ____: | ____: | ____: false |
| fast | ____: | ____: | ____: | ____: | ____: | ____: | ____: slow |
| hard | ____: | ____: | ____: | ____: | ____: | ____: | ____: soft |
| ugly | ____: | ____: | ____: | ____: | ____: | ____: | ____: beautiful |
| wise | ____: | ____: | ____: | ____: | ____: | ____: | ____: foolish |
| masculine | ____: | ____: | ____: | ____: | ____: | ____: | ____: feminine |

Each subject was given a standard set of semantic differential instructions on a cyclostyled sheet (shown in Appendix I) which was adapted from Osgood et al. (1957). The experimenter read through the instructions with the subject, who was required to "check" the form in the appropriate manner in order to learn how the scales were to be "checked" on the actual test form.

THE PAIRED COMPARISONS

Thirteen concepts were selected from a previous scaling study by Arthur (1965c) using ratings on a standard semantic differential. On the basis of this study, the chosen concepts had known scale values distributed over the entire range of the "good-bad" dimension, with a fairly even spread. Table 2 gives the concepts with their scale values.

TABLE 2

Scale values of concepts on a "good-bad" dimension of the Semantic Differential (Arthur 1965c)

| Concept | Scale Value | Concept | Scale Value |
|---|---|---|---|
| PEACE | 0.225 | DIRT | 4.236 |
| FATHER | 0.912 | DANGER | 4.437 |
| BATH | 1.512 | PAIN | 4.639 |
| LEMON | 2.211 | ABORTION | 4.854 |
| WINTER | 3.129 | DEVIL | 5.205 |
| DARK | 3.603 | TORTURE | 5.590 |
| FLEA | 3.984 | | |

The thirteen concepts were paired in all possible combinations, including the position of the concept in the pair, thus giving a total of 156 comparisons in all,

that is, 24 judgements on each of the thirteen concepts. The comparisons were made by means of cards on each of which a pair of concepts were typed in capitals. In addition, the cards were ordered so that the concepts in any given pair did not occur in the previous or the following pairs. The order of presentation was the same for all subjects. The verbal responses of each subject were recorded by the examiner on a prepared data sheet as shown in Appendix II. Each subject was required to make judgements on all the concepts on two separate occasions, in the one case the judgement was to be made in terms of which was "better", and on the other which was "worse". Subjects were randomly assigned to two groups, group 1 making the "better" judgements first, and group 2 making the judgement "worse" first. All subjects were given a typed form with the following instructions which were read aloud by the examiner:

> "In this test you are required to compare pairs of words. First think of what each word in the pair means to you, and then say which one you consider 'better'. Now look at the first pair of words, 'PEACE' and 'FLEA'. Think of what each means to you: now say which you consider 'better'

If you are not sure, guess".

(Note: 'worse' was substituted for 'better' when judgements of 'worse' were to be made).

DESIGN

Subjects.

Forty-four unselected patients at a large mental hospital were tested, the sample being made up of 19 males and 25 females. Old, i.e. over 50 years, organic and mentally defective patients were excluded.

Controlled Variables.

Measures were taken of age and intelligence. Age in years was given by the patient. Intelligence was measured on the Wechsler Adult Intelligence Scale Vocabulary sub-test, the raw score only being used. The vocabulary was thought to be the most appropriate measure considering the verbal nature of the experimental task. (Note: Tables giving the mean age and intelligence of the experimental groups are given under "Results" since the composition of the groups was not determined until their responses in the Semantic Differential test had been analysed).

Order of Testing.

Every patient was given the W.A.I.S. vocabulary first. This was done so that the mental defectives could be eliminated at an early stage. No subjects were in fact eliminated. This was followed by the Paired Comparisons judgements in terms of 'better' or 'worse' depending on whether the patient had been assigned to group 1 or group 2. The Semantic Differential was then administered, and the testing was completed with the Paired Comparisons judging 'worse' or 'better'. The full testing required from 60 to 90 minutes for each patient.

# CHAPTER 4

## RESULTS

### DIFFERENTIATION OF THE EXPERIMENTAL GROUPS

The responses of the 44 subjects tested on the Semantic Differential were summed into individual frequency scores on categories 1 and 7 (extreme), 2, 3, 5 and 6 (non-extreme), and 4 (neutral). See Appendix III for the individual frequencies. Figure 1 shows the distribution of frequencies for extreme, neutral and non-extreme positions. This graph did not show any clear bimodal distributions and there was no certainty that subjects with a moderately high frequency in the extreme categories would not also have a considerable frequency in the non-extreme categories. To obtain this condition it was decided to determine cut-off points which would produce a maximum number of subjects in each group with no overlap between the groups by plotting the frequency of response in the extreme category against frequency in the non-extreme category as shown in Figure 2. By inspection, the best cut-offs were found to be as follows:

Extreme Group: 35 and above on categories 1 and 7.
below 20 on categories 2, 3, 5 and 6.

Figure 1.
Frequency Distribution
of Scale Position Usage.

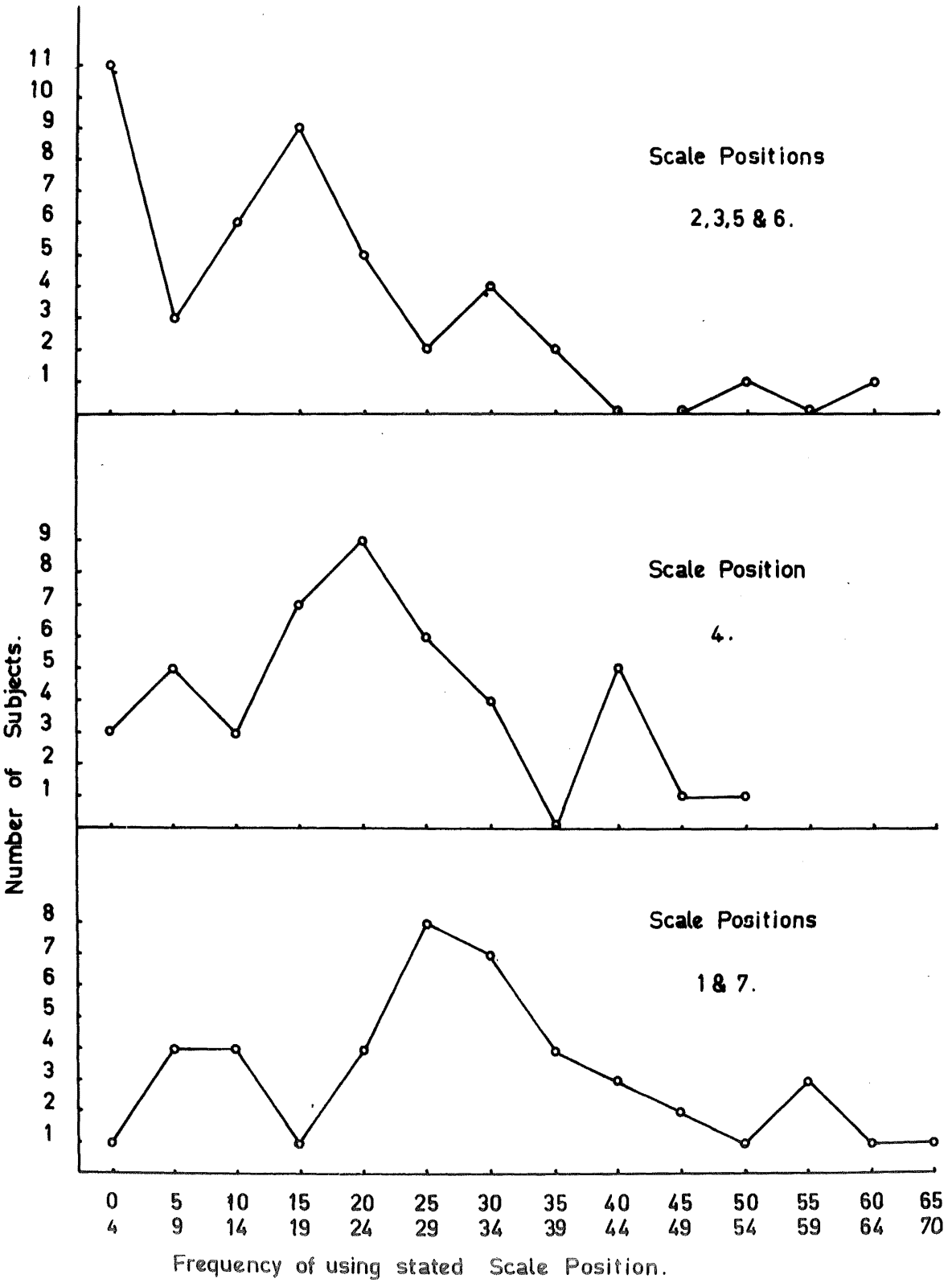
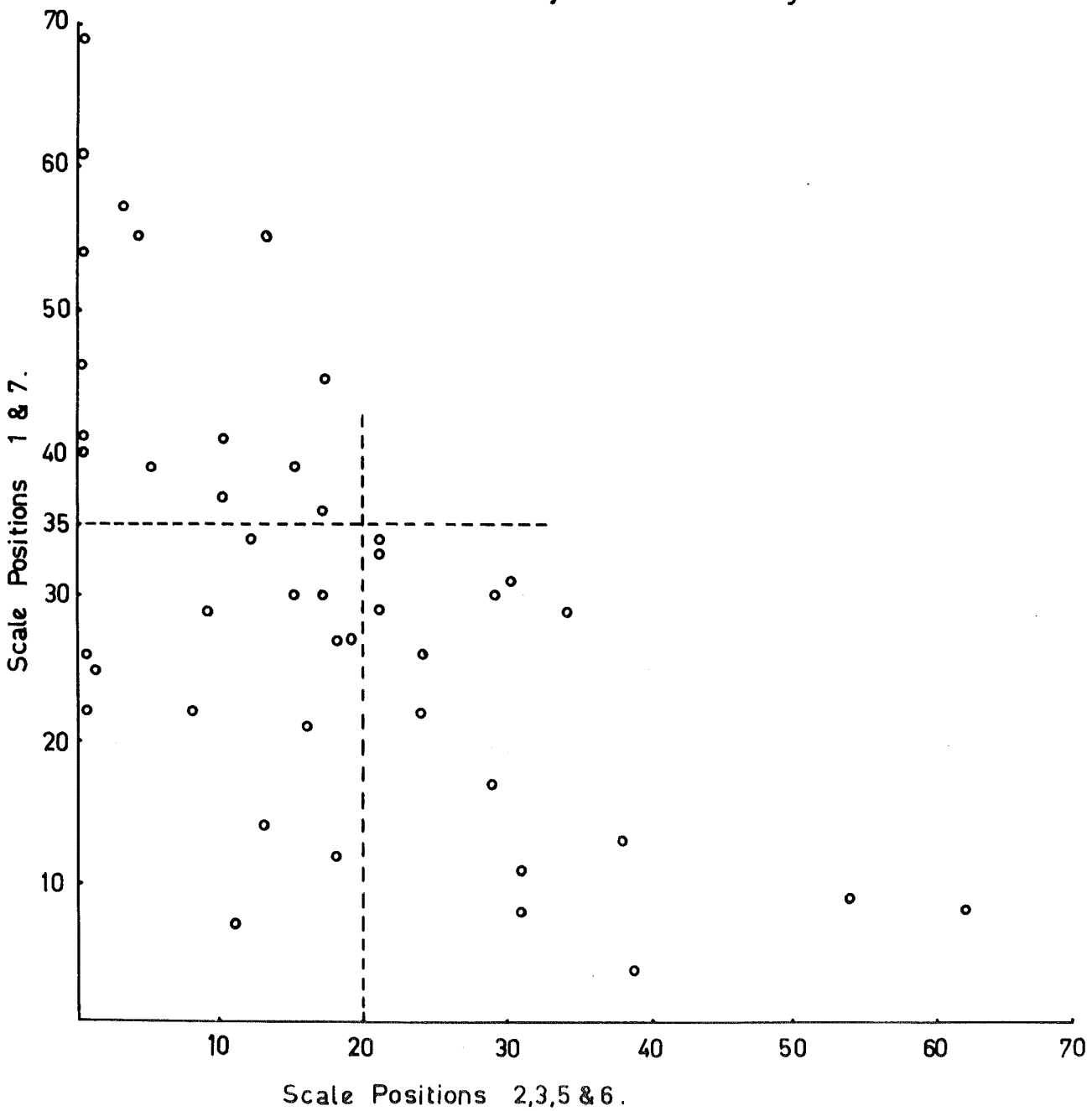


Figure 2.
Frequency of Using
Extreme (1&7), and Non-extreme (2,3,5&6)
Scale Positions by Individual Subjects.



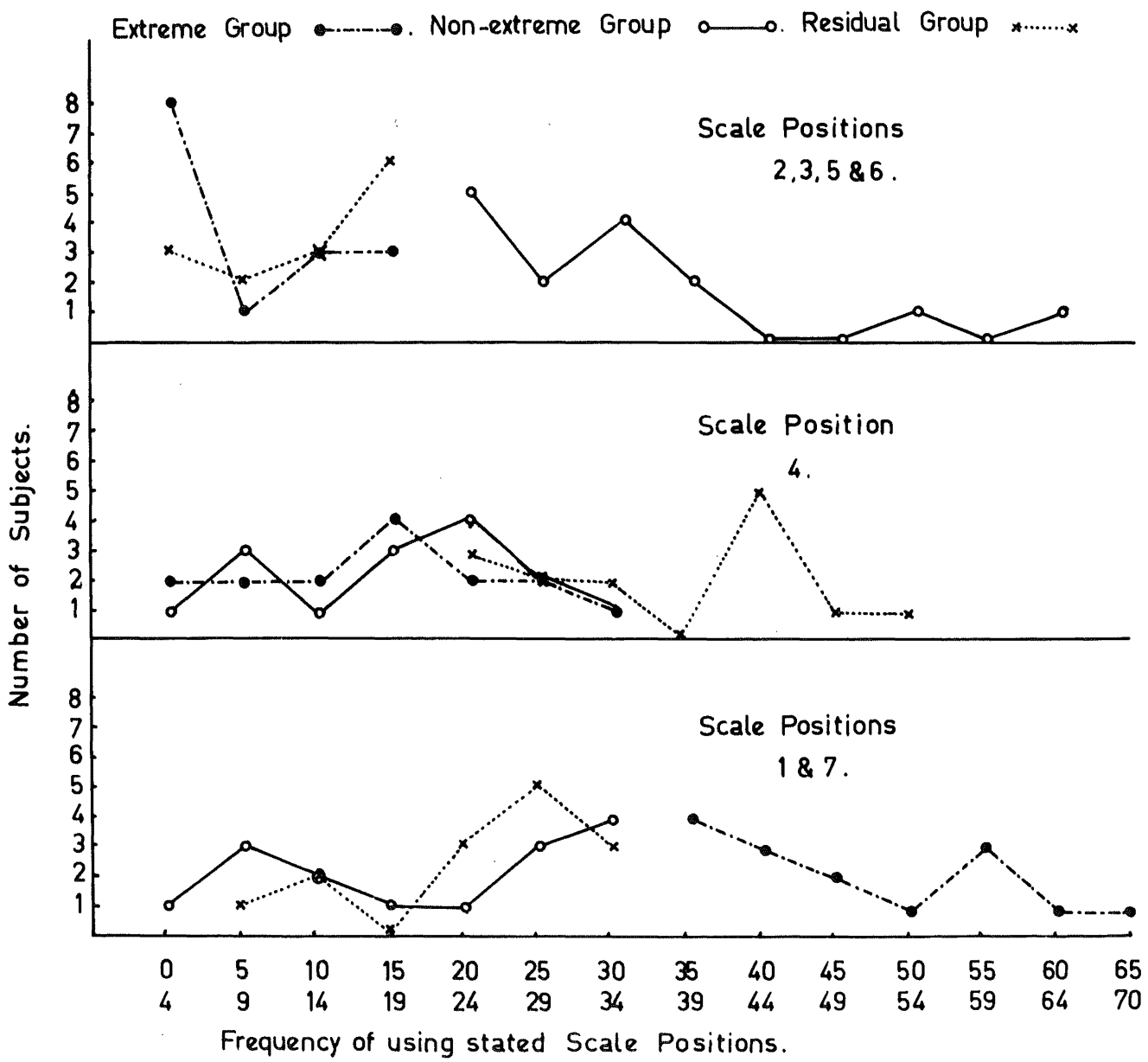
Non-extreme Group: 20 and above on categories, 2,3, 5 and 6.

below 35 on categories 1 and 7.

The use of these two cut-off points divided the subjects into three distinct groups of approximately equal numbers. It is of course arbitrary to divide the subjects into groups in terms of extremity of response, thus it was necessary to ascertain whether those groups identified by frequency usage of the extreme and non-extreme categories, also differed in respect of their use of the neutral category. Figure 3 shows complete overlap between the groups in checking the neutral position 4, and no overlap in checking the other positions. This meant that the differences between the two groups were accounted for in their use of the extreme and non-extreme positions.

It is to be noted that the residual group (neither extreme nor non-extreme) is not within the focus of interest of the present study and thus no analysis was carried out on their responses. Nevertheless, the smallness of this group emphasises the fact that extreme and non-extreme raters are not some minor ends of some normal distribution of subjects but are major

Figure 3.
Composition of Groups by Frequency
Distribution of Scale Position Usage.



parts of the distribution.

A summary of the composition of the three groups by age, sex, intelligence (raw scores) and frequency of response in the extreme, non-extreme and neutral categories can be found in Table 3. A 't' test was conducted on the Extreme and Non-extreme groups and the results of this have been included in Table 3. The only significant differences between the two groups were those on the usage frequencies of the extreme and non-extreme categories.

Composition of groups by mean age, sex, intelligence and frequency of scale position usage on the Semantic Differential

| | Non-extreme | | Extreme | | Residual | | Non-extreme |
|---|---|---|---|---|---|---|---|
| MALE | 8 | | 5 | | 6 | | v Extreme |
| FEMALE | 7 | | 10 | | 8 | | 't' Test |
| | X | S.D. | X | S.D. | X | S.D. | |
| AGE | 27.0 | 7.99 | 32.8 | 9.51 | 30.0 | 8.50 | 0.117 |
| INTELLIGENCE | 50.3 | 14.60 | 49.0 | 14.19 | 48.1 | 13.15 | 0.812 |
| FREQUENCY 1 & 7 | 20.3 | 10.29 | 47.6 | 9.78 | 23.3 | 5.35 | 7.280 * |
| FREQUENCY 4 | 16.9 | 8.63 | 16.1 | 8.83 | 35.5 | 9.71 | 0.073 |
| FREQUENCY 2, 3, 5 & 6 | 32.5 | 11.50 | 6.3 | 6.53 | 11.2 | 6.55 | 7.440 * |

* $p < .001$

SCALING

The scaling was carried out under the assumptions of Case V for incomplete data on a paired comparison (Edwards 1957a). Appendix IV shows the raw data from which the scaling was done. The scale values for the two groups and two judgement conditions are in Table 4. Figure 4 shows these scale values plotted against a unit rank order taken from the non-extreme group judging 'better'. As can be seen there are large differences in the width of scale between the two groups both in their judgements in terms of 'better' and in terms of 'worse'. A Mann-Whitney 'U' Test (Siegel 1956) was carried out on the differences between the scale values. The differences between the Extreme and Non-extreme groups for both 'better' and 'worse' judging conditions were found to be significant (see Table 5). Thus the hypothesis that the Extreme and Non-extreme groups would differ in their scaling on an evaluative dimension where extreme judgements are not possible is confirmed.

Table 5 also shows that there is no significant difference within the two groups over the two judging conditions, 'worse' and 'better'. This confirms the

TABLE 4

Scale values from paired comparisons

| Concept | Better | | Worse* | |
|---|---|---|---|---|
| | Non-extreme | Extreme | Non-extreme | Extreme |
| PEACE | 0.000 | 0.000 | 0.000 | 0.003 |
| FATHER | 0.849 | 0.041 | 1.141 | 0.000 |
| BATH | 0.868 | 0.191 | 1.228 | 0.224 |
| LEMON | 1.178 | 0.744 | 1.485 | 0.585 |
| WINTER | 1.936 | 0.772 | 2.287 | 1.070 |
| DARK | 2.148 | 0.788 | 2.392 | 0.991 |
| FLEA | 2.419 | 1.217 | 2.764 | 1.261 |
| DIRT | 2.725 | 1.323 | 3.142 | 1.432 |
| PAIN | 3.096 | 1.431 | 3.452 | 1.570 |
| DANGER | 3.198 | 1.391 | 3.756 | 1.647 |
| DEVIL | 3.808 | 1.985 | 4.096 | 1.965 |
| TORTURE | 3.966 | 1.728 | 4.461 | 1.968 |
| ABORTION | 4.008 | 1.910 | 4.408 | 2.103 |

(* scale positions are reversed)

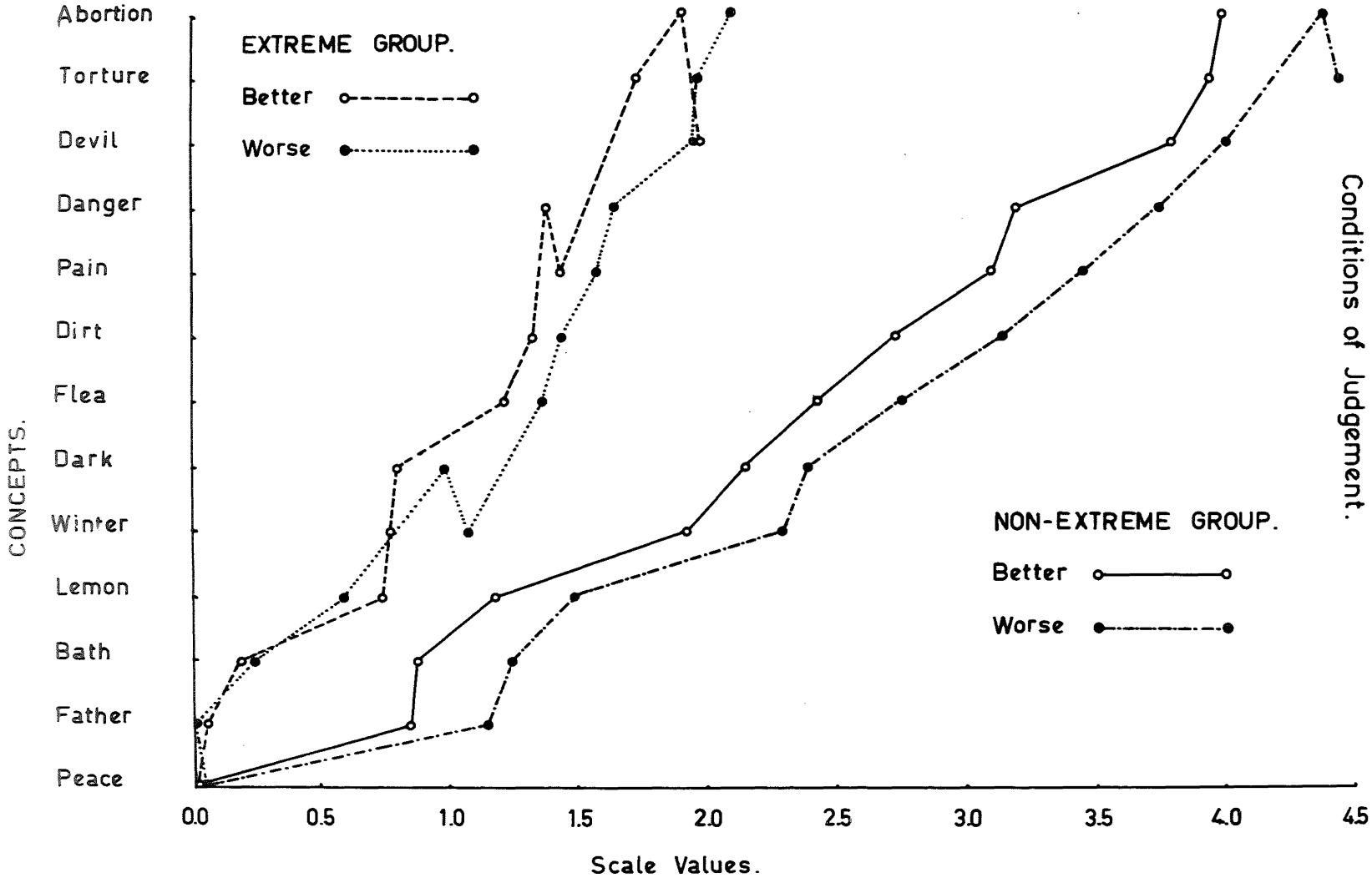



Figure 4.
Scale Values of the Extreme & Non-extreme Groups for Both Conditions of Judgement.

hypothesis that the scaling of the 'worse' judgements from a given group when reversed would not differ significantly from the 'better' judgements for the same group.

Since the major interest was in measuring mean differences in the scaling behaviour of the two groups, and not in preserving the individual identities of the scale values, it was decided to use the Mann-Whitney 'U' test as the most appropriate method of determining the significance of such differences. Table 5 shows that there were statistically significant differences in the scaling behaviour of the Extreme and Non-extreme groups under both better and worse judging conditions. On the other hand, the differences between better and worse judgements of each group were found to be not significant.

Using the Wilcoxon test (Siegel, 1956) which preserves the individual identities of the scales, it was found that the difference between better and worse judgements for both the Extreme and Non-extreme groups was statistically significant. The disparity in the results from these two non-parametric tests suggests that the difference being measured by Wilcoxon, although real in a statistical sense, is not of sufficient magnitude to be psychologically useful or clinically useful and it has thus been disregarded.

TABLE 5

Mann-Whitney 'U' test on the differences between the scale values

| | 'U' | signif. |
|---|---|---|
| Extreme Better v Non-extreme Better | 24.5 | $p < .002$ |
| Extreme Worse v Non-extreme Worse | 32.5 | $p < .02$ |
| Extreme Better v Extreme Worse | 95.5 | N.S. |
| Non-extreme Better v Non-Extreme Worse | 99.5 | N.S. |

hypothesis that the scaling of the 'worse' judgements from a given group when reversed would not differ significantly from the 'better' judgements for the same group.

ANALYSIS OF SCALING DIFFERENCES

Having found a significant difference between the scale values of the Extreme and Non-extreme groups it was necessary to explore the possible reasons or explanations for this difference. It was noted from their scale values, that for the Extreme group the concepts used were much closer together than they were for the Non-extreme group. This narrowing of the scales obtained from Paired Comparisons can only occur as a result of increased dispersion of judgements. The increased dispersion may be due to the judgements being more inconsistent in the Extreme group in at least two ways: first, the individual judge may be inconsistent, that is, his value ordering is inconsistent. Secondly, the individual while being highly consistent in himself, does not agree with the rest of the group he is in, that is, he is not consistent with the group as a whole.

To measure the first type of inconsistency it was necessary to determine the degree to which the paired comparison choices of each judge were consistent with his simple rank ordering of the concepts, and then to see whether the Extreme group were as consistent in

this respect as were the Non-extreme group. Kendall (1948) has shown that if a subject expresses preferences for three objects X, Y, Z, as $X<Y<Z<X$ then it can be said that the triad X, Y, Z, is 'circular' or 'inconsistent'. The more circular triads there are, the further the departure from a ranking situation toward a position of inconsistency where ranking does not occur. Table 6 shows the number of circular triads for individual subjects of the two groups under both judging conditions. A Mann-Whitney 'U' Test was used to test whether the two independent groups have been drawn from the same population. The result shown in Table 6 indicated that there was no significant difference between the groups in the degree of inconsistency either for judgements of 'better' or for 'worse'. The raw data for the calculation of the circular triads is shown in Appendix V.

The second type of inconsistency, that is the homogeneity of the groups, was measured by taking the means and standard deviations of the frequency with which each concept was judged as 'better' and as 'worse' for each group and these are summarised in Table 7. The raw data for this calculation is shown in Appendix

TABLE 6

Number of circular triads for individual subjects and Mann-Whitney 'U' values for comparison of the two groups

| Better | | Worse | |
|---|---|---|---|
| Non-extreme | Extreme | Non-extreme | Extreme |
| 52 | 11 | 20 | 8 |
| 5 | 7 | 18 | 2 |
| 4 | 6 | 10 | 3 |
| 2 | 3 | 4 | 1 |
| 5 | 7 | 2 | 4 |
| 5 | 3 | 7 | 8 |
| 4 | 2 | 5 | 2 |
| 13 | 24 | 6 | 31 |
| 2 | 3 | 4 | 6 |
| 0 | 3 | 2 | 3 |
| 4 | 18 | 9 | 21 |
| 6 | 21 | 8 | 25 |
| 2 | 20 | 8 | 8 |
| 2 | 13 | 1 | 3 |
| 0 | 3 | 3 | 19 |
| 'U' 70 | N.S. | 'U' 108 | N.S. |

TABLE 7

Means and standard deviations of judgements 'better' and 'worse' for each concept by each group

| Concept | 'Better' | | | | 'Worse'* | | | |
|---|---|---|---|---|---|---|---|---|
| | Non-extreme | | Extreme | | Non-extreme | | Extreme | |
| | X | S.D. | X | S.D. | X | S.D. | X | S.D. |
| PEACE | 23.26 | 1.10 | 20.40 | 5.90 | 23.60 | 0.63 | 20.73 | 5.22 |
| FATHER | 20.20 | 3.32 | 20.13 | 5.42 | 20.13 | 3.42 | 20.86 | 5.13 |
| BATH | 19.73 | 1.22 | 18.46 | 4.79 | 19.40 | 1.45 | 18.86 | 3.89 |
| LEMON | 18.33 | 1.35 | 14.66 | 5.01 | 18.53 | 1.55 | 16.60 | 5.48 |
| WINTER | 14.73 | 1.19 | 14.53 | 3.82 | 14.53 | 1.77 | 12.73 | 2.96 |
| DARK | 13.86 | 2.39 | 14.13 | 4.12 | 14.40 | 3.29 | 13.46 | 3.50 |
| FLEA | 12.00 | 2.80 | 10.73 | 4.30 | 11.73 | 3.03 | 10.80 | 3.84 |
| DIRT | 9.93 | 3.30 | 9.80 | 2.76 | 10.00 | 2.10 | 10.46 | 3.56 |
| PAIN | 7.93 | 2.31 | 8.60 | 4.79 | 8.93 | 2.09 | 8.80 | 5.00 |
| DANGER | 7.13 | 2.87 | 9.13 | 4.95 | 6.26 | 1.67 | 8.00 | 4.28 |
| DEVIL | 3.53 | 2.75 | 4.33 | 3.54 | 3.86 | 3.40 | 5.40 | 4.79 |
| TORTURE | 2.86 | 2.26 | 6.20 | 6.11 | 2.46 | 1.73 | 5.13 | 5.68 |
| ABORTION | 2.46 | 2.82 | 4.86 | 6.13 | 2.80 | 2.37 | 4.13 | 4.58 |

*scores are reflected

V. A low standard deviation suggests a homogeneous group and the standard deviation increases as the group gets less homogeneous or more inconsistent. A Mann-Whitney 'U' Test on the distributions of means and standard deviations (see Table 8) showed that while the distributions of the means were not significantly different, the distributions of the standard deviations for the Extreme and Non-extreme groups were highly significant, the Extreme group having the larger standard deviations and thus being the more heterogeneous group.

TABLE 8

Mann-Whitney 'U' Test on the differences between distributions of means and standard deviations (S.D.)

| | 'U' Means | 'U' Signif. | 'U' S.D. | 'U' Signif. |
|---|---|---|---|---|
| Extreme Better v Non-extreme Better | 84 | N.S. | 70 | N.S. |
| Extreme Worse v Non-extreme Worse | 85 | N.S. | 94 | N.S. |
| Extreme Better v Extreme Worse | 82 | N.S. | 4 | $p < .002$ |
| Non-extreme Better v Non-Extreme Worse | 83 | N.S. | 4 | $p < .002$ |

TABLE 8a

Wilcoxon Matched-Pairs Signed-Ranks Test on the within group differences between distributions of means and standard deviations

| | MEANS | | | S.D. | | |
|---|---|---|---|---|---|---|
| | T | N | Signif | T | N | Signif |
| Extreme Better v Extreme Worse | 46 | 13 | N.S. | 28 | 13 | N.S. |
| Non-Extreme Better v Non-extreme Worse | 38.5 | 13 | N.S. | 45 | 13 | N.S. |

## CHAPTER 5

### DISCUSSION OF RESULTS

The first aim of this study was to show that psychiatric patients can be differentiated into two groups in terms of extreme responses with the extreme responders falling naturally into a different category from the rest in respect of checking extreme positions but not of the neutral position. The frequency usage of the extreme positions on the Semantic Differential was thus expected to give a bimodal distribution. However the results show that there is no such bimodality, rather the obtained curve is multimodal necessitating the use of a rather arbitrary method of selecting cut-off points for differentiating the groups. Nevertheless the division used, (that of choosing a cut-off point which placed maximum numbers in each group with no overlap), seems to make a natural distinction since the groups have been shown not to differ in their use of the neutral category.

It appears from the results that there are actually three types of raters, Extreme, Non-extreme and Neutral. The present study was not aimed at demonstrating the existence of a neutral group, neither has it investigated the judgemental behaviour of such a group.

However it does suggest that this group should be taken into account in future studies, the expectation being that in number it would comprise one third of any unselected sample from a psychiatric population, the other two-thirds being equally spread between the extreme and non-extreme raters.

The major interest of this study was to investigate the possibility that the extreme response bias is attributable at least in part to a difference in the cognitive judgement process. It was postulated that such a difference would show itself when the scaled comparative judgements of a group of extreme responders were compared with the judgements of a group of non-extreme responders. The results indicate that there is a clear difference in the scale widths of the Extreme and Non-extreme groups for judging as 'better' ($p < .002$) and 'worse' ($p < .02$). However a further analysis of the nature of this difference reveals that while both groups are highly consistent in their judgements, the responses of the Extreme group are more heterogeneous, and in consequence the width of scale is narrowed for this group. The difference in scaling could be explained entirely by this variation in the homogeneity of

response, and therefore the study does not throw any light on the nature of the extreme response bias from the point of view of cognitive variables like judgement.

The main contribution of the present study is in demonstrating that extreme raters are more heterogeneous in the meanings they attribute to ordinary concepts. This is not to say that the subjects are different, but they do give a greater variety of meaning to certain concepts. The scaling approach used in the present study is sensitive to even minor shifts in meaning, especially where small groups are used. In the present study the groups were small enough (only 15 in each) for one or two subjects rating with a reversed meaning polarity on some of the words to have a great effect on the total dispersion, and actually collapse the width of the scale.

Appendix V shows the frequency of judging the concepts as 'better' and 'worse' for the Extreme and Non-extreme groups. In each case the concepts are placed in rank order according to scaled values from the given group and judging condition. By inspection of these tables it can be seen that subjects Nos. 14 and 15 in the Extreme group both reversed the meaning polarity of the concepts on a considerable number of times. Other

subjects in the Extreme group show a similar trend to a lesser degree.

It is interesting to note that the two subjects 14 and 15 who gave the most reversals have both done acts of self mutilation and made numerous attempts at suicide. They considered 'better' such concepts as 'torture', 'pain' and 'abortion', while 'bath', 'father' and 'peace' were considered as 'worse'. The fact that they reversed the polarity of those concepts which appear relevant to their symptoms, indicates that meaningfulness may be an important variable in the extreme response. Extreme responses may reflect high intensity of meaning. A few other subjects also tended to reverse the polarity of such concepts as 'danger', 'devil' and 'torture'.

Further research is indicated in two areas. First is to investigate the judgemental behaviour of the 'Neutral' group found in this study, and secondly to explore the effects of the meaningfulness of concepts related to a given symptomatology and extreme response by comparing the responses of normals with patients having the given symptoms.

# CHAPTER 6

## SUMMARY AND CONCLUSIONS

There has been considerable evidence in the literature for the existence of extreme response bias and for its correlates with pathology and personality. The present study however sets out to investigate the psychological nature of the extreme response by having extreme and non-extreme responders make judgements on a task which does not allow extreme responses.

The aim was first to distinguish groups of extreme and non-extreme responders by using a task which permitted graded responses in terms of intensity (the Semantic Differential), and then to measure the evaluative judgement continua by scaling the responses of these two groups from a task where extreme responses are not possible (the Paired Comparisons). The hypothesis was that since extreme and non-extreme responders differ in their graded responses they may also differ in the behaviour which is measured by scaling.

The extreme and non-extreme groups were arbitrarily distinguished by the frequency with which they used the extreme and non-extreme categories in the Semantic Differential. It was found that three distinct groups were obtained, the third group having their

responses mainly in the neutral category.

Case V scaling procedure of the Paired Comparisons judgements revealed a difference in the width of scales the extreme group having a narrower scale than the non-extreme group. Dispersion of responses was studied and the extreme group was found to be the more heterogeneous in this respect. Since the scaling difference can be most satisfactorily explained in terms of the greater heterogeneity of response in the extreme group, it must be concluded that the difference does not throw any light on the nature of the extreme response.

It was observed that there was a reversal in the polarity of some of the concepts by a number of extreme raters, and this seemed to be related to their symptoms. Further research in this area was recommended.

## REFERENCES


ARTHUR, A.Z. (1965a) Explanation of response bias. An unpublished manuscript.

ARTHUR, A.Z. (1965b) An unpublished manuscript.

ARTHUR, A.Z. (1965c) An unpublished manuscript.

ARTHUR, A.Z. (1966) Response bias in the Semantic Differential. (To appear in Brit. J. clin. soc Psychol.)

BARNES, E.H. (1955) The relationship of biased test responses to pathology. J. abnorm. soc. Psychol., 51, 286-290.

BERG, I.A. (1955) Response bias and personality: the deviation hypothesis. J. psychol., 40, 61-72.

BERG, I.A. (1961) Measuring deviant behaviour by means of deviant response sets. In I.A. Berg and B.M. Bass (Eds), Conformity and deviation. New York: Harper.

BERG, I.A., COLLIER, J.S. (1953) Personality and group differences in extreme response sets. Educ. Psychol. Measmt., 13, 164-169.

BORGATTA, E.F., GLASS, D.C. (1961) Personality concomitants of extreme response sets. (ERS). J. abnorm. soc. Psychol., 55, 213-221.

COUCH, A., KENISTON, K. (1960) Yeasayers and naysayers: agreeing response set as a personality variable. J. abnorm. soc. Psychol., 60, 151-174.

CRONBACH, L.J. (1950) Further evidence on response sets and test design. Educ. Psychol. Measmt., 10, 3-31.

CRONBACH, L.J., GLESER, G.C. (1953) Assessing similarity between profiles. Psychol. Bull., 50, 456-473.

DONAHOE, J.W. (1961) Changes in meaning as a function of age. J. genet. Psychol., 99, 23-28.

EDWARDS, A.L. (1957a) Techniques of attitude scale construction. New York: Appleton-Century-Crofts, Inc.

EDWARDS, A.L. (1957b) The social desirability variable in personality research. New York: Dryden.

EDWARDS, A.L. (1959) Social desirability and personality test construction. In B.M. Bass and I.A. Berg (Eds). Objective approaches to personality assessment. Princeton, N.J.: Van Nostrand.

JACKSON, D.N., MESSICK, S. (1958) Content and style in personality assessment. Psychol. Bull., 55, 243-252.

JOURARD, S.M. (1959) Self-disclosure and other-cathexis. J. abnorm. soc. Psychol., 59, 428-431.

KELLEY, H.H., HOVLAND, C.I., SCHWARTZ, M., ABELSON, R.P. (1955) The influence of judges attitudes in the methods of attitude scaling. J. Soc. Psychol., 42, 147-158.

KENDALL, M.G. (1948) Rank correlation methods. London: Charles Griffin & Co. Ltd.

LEWIS, N.A., TAYLOR, J.A. (1955) Anxiety and extreme response preferences. Educ. Psychol. Measmt., 15, 111-116.

McGEE, R.K. (1962) Response style as a personality variable. Psychol. Bull., 59, 284-295.

NEURINGER, C. (1963) Effects of intellectual level and neuropsychiatric status on the diversity of intensity of semantic differential ratings. J. consult. Psychol., 27, 280.

O'DONOVAN, D. (1965) Rating extremity: pathology or meaningfulness? Psychol. Rev., 72, 358-372.

OSGOOD, C.E., SUCI, G.J., TANNENBAUM, P.H. (1957). The measurement of meaning. Urbana, Ill: University of Illinois Press.

PARSONSON, B.S. (1965) Differentiation of psychiatric groups by means of subjective meanings of symptoms, unpublished Masters thesis. University of Canterbury, New Zealand.

RORER, L.G. (1965) The great response-style myth. Psychol. Bull., 63, 129-156.

SIEGEL, S. (1956) Non-parametric statistics: for the behavioural sciences. New York: McGraw-Hill Book Co. Inc.

SMITH, E.E. (1959) Defensiveness, insight and the 'K' scale. J. consult. Psychol., 60, 151-174.

WARE, E.E. (1958) In A.Z. Arthur (1965a).

WERTHEIMER, R., McKINNEY, F. (1952) A case history blank as a projective technique. J. consult. Psychol., 16, 49-60.

ZAX, M., GARDINER, D.W., LOWRY, D.G. (1964) Extreme response tendency as a function of emotional adjustment. J. abnorm. soc. Psychol., 69, 654-657.

# APPENDICES I-V

APPENDIX I

Instructions for the Semantic Differential

APPENDIX II

Data recording sheet for the Paired Comparisons

APPENDIX III

Subjects by Age, Sex, Intelligence and Frequency of Scale Position Usage on the Semantic Differential

APPENDIX IV

Frequency of 'better' and 'worse' Paired Comparisons responses on all concepts for Extreme and Non-extreme groups

APPENDIX V

Frequency of 'better' and 'worse' Paired Comparisons responses for individual subjects

APPENDIX I

## INSTRUCTIONS FOR THE SEMANTIC DIFFERENTIAL

The purpose of this test is to see what certain things mean to various people by having them judge them against descriptive adjectives. In taking this test please make your judgements according to what these things mean to you. On each page of this booklet you will find a different thing to be judged and beneath it a number of lines with adjectives to judge against. You are to judge the thing on each of the lines. Here is how you judge the lines.

If you feel that the thing at the top of the page is very like one end of the line, you should put your check-mark like this:

slow : ____ : ____ : ____ : ____ : ____ : ____ : ____ : fast

or

slow : ____ : ____ : ____ : ____ : ____ : ____ : ____ : fast

If you feel that the thing is quite like one end of the line (but not extremely) you should put your check-mark like this:

strong : ____ : ____ : ____ : ____ : ____ : ____ : ____ : weak

or

strong : ____ : ____ : ____ : ____ : ____ : ____ : ____ : weak

If the thing seems only slightly like one end as opposed to the other (but not really neutral), then you should check like this:

good : ____ : ____ : ____ : ____ : ____ : ____ : ____ : bad

or

good : ____ : ____ : ____ : ____ : ____ : ____ : ____ : bad

The end toward which you check, of course, depends on which of the two ends of the line seem better to describe the thing you're judging.

If you think the thing you are judging to be neutral as regards the adjectives, or if the adjectives are completely irrelevant, not related to the thing, then you should put your check-mark in the middle:

kind : ____ : ____ : ____ : ____ : ____ : ____ : ____ : cruel

Important

1. Be sure you check every line of each page of the booklet do not omit any

2. Never put more than one check-mark on a single line.

Work quickly through the test. Do not worry or puzzle over individual items, it is your first impressions, the immediate "feelings" about the items, that we want. On the other hand, please do not be careless because we want your true impressions.

# APPENDIX II

1 Peace - Flea
2 Abortion - Winter
3 Lemon - Pain
4 Winter - Danger
5 Devil - Torture
6 Bath - Peace
7 Torture - Father
8 Dirt - Devil
9 Pain - Father
10 Dark - Devil
11 Flea - Bath
12 Father - Abortion
13 Danger - Bath
14 Peace - Lemon
15 Abortion - Bath
16 Winter - Dark
17 Bath - Torture
18 Devil - Winter
19 Torture - Flea
20 Dirt - Father
21 Pain - Flea
22 Dark - Dirt
23 Flea - Danger
24 Danger - Dark
25 Father - Bath
26 Peace - Pain
27 Abortion - Danger
28 Lemon - Torture
29 Winter - Devil
30 Torture - Lemon
31 Devil - Abortion
32 Bath - Winter
33 Dirt - Flea
34 Pain - Lemon
35 Dark - Father
36 Danger - Devil
37 **Flea - Dark**
38 Father - Danger
39 Peace - Torture
40 Abortion - Dark
41 Lemon - Winter
42 Devil - Bath
43 Winter - Dirt
44 Bath - Abortion
45 Torture - Pain
46 Dirt - Lemon
47 Pain - Peace
48 Dark - Flea
49 Danger - Dirt
50 Flea - Devil
51 Father - Dark
52 Peace - Winter
53 Abortion - Devil
54 Winter - Father
55 Lemon - Abortion
56 Devil - Danger
57 Torture - Peace
58 Bath - Danger
59 Dirt - Pain
60 Dark - Lemon
61 Pain - Torture
62 Flea - Dirt
63 Danger - Father
64 Peace - Abortion
65 Father - Devil
66 Abortion - Dirt
67 Lemon - Bath
68 Winter - Flea
69 Devil - Dark
70 Torture - Winter
71 Bath - Dark
72 Dirt - Peace
73 Pain - Winter
74 Flea - Father
75 Dark - Pain
76 Danger - Flea
77 Father - Dirt
78 Peace - Bath
79 Abortion - Father
80 Lemon - Danger
81 Devil - Dirt
82 Winter - Lemon
83 Bath - Devil
84 Torture - Abortion
85 Lemon - Peace
86 Dirt - Torture
87 Pain - Abortion
88 Dark - Peace
89 Flea - Lemon
90 Peace - Danger
91 Abortion - Flea
92 Danger - Lemon
93 Father - Flea
94 Lemon - Dark
95 Winter - Pain
96 Devil - Father
97 Bath - Dirt
98 Dark - Torture
99 Pain - Bath
100 Dirt - Winter
101 Torture - Bath
102 Flea - Pain
103 Father - Lemon
104 Danger - Pain
105 Peace - Dark
106 Abortion - Lemon
107 Lemon - Devil
108 Winter - Peace
109 Devil - Flea
110 Bath - Father
111 Torture - Danger
112 Dirt - Abortion
113 Pain - Danger
114 Dark - Winter
115 Flea - Peace
116 Father - Pain
117 Danger - Peace
118 Abortion - Pain
119 Peace - Devil
120 Lemon - Dirt
121 Winter - Torture
122 Devil - Lemon
123 Bath - Flea
124 Torture - Dark
125 Dirt - Bath
126 Pain - Dark
127 Flea - Torture
128 Dark - Abortion
129 Danger - Torture
130 Father - Peace
131 Winter - Abortion
132 Peace - Dirt
133 Lemon - Father
134 Abortion - Peace
135 Devil - Pain
136 Bath - Lemon
137 Torture - Devil
138 Dirt - Dark
139 Pain - Devil
140 Dark - Bath
141 Flea - Winter
142 Father - Torture
143 Danger - Winter
144 Peace - Father
145 Abortion - Torture
146 Lemon - Flea
147 Winter - Bath
148 Devil - Peace
149 Bath - Pain
150 Torture - Dirt
151 Dark - Danger
152 Pain - Dirt
153 Flea - Abortion
154 Dirt - Danger
155 Father - Winter
156 Danger - Abortion

# APPENDIX III

Individual subjects by age, intelligence (vocabulary raw score), sex, and frequency of scale position usage on the Semantic Differential

## EXTREME GROUP

| Subject No. | 1 | 2 | 3 | 4 | 5 | 6 | 12 | 14 | 15 | 17 | 18 | 19 | 20 | 41 | 43 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Age | 15 | 22 | 38 | 19 | 29 | 39 | 39 | 21 | 27 | 45 | 38 | 36 | 39 | 45 | 41 |
| Intelligence | 45 | 29 | 47 | 54 | 52 | 68 | 59 | 67 | 52 | 48 | 74 | 21 | 39 | 33 | 46 |
| Sex | M | F | F | F | M | M | M | F | F | F | F | F | F | F | I |
| Frequency of scale position usage | | | | | | | | | | | | | | | |
| 1 and 7 | 57 | 40 | 41 | 37 | 41 | 46 | 36 | 54 | 55 | 55 | 39 | 69 | 61 | 39 | 4 |
| 4 | 10 | 30 | 29 | 23 | 19 | 24 | 17 | 16 | 2 | 11 | 26 | 1 | 9 | 16 | |
| 2, 3, 5 and 6 | 3 | 0 | 0 | 10 | 10 | 0 | 17 | 0 | 13 | 4 | 5 | 0 | 0 | 15 | 1 |

# APPENDIX III

## NON-EXTREME GROUP

| Subject No. | 7 | 10 | 11 | 13 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 37 | 38 | 42 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Age | 29 | 31 | 29 | 18 | 29 | 27 | 40 | 17 | 28 | 15 | 33 | 18 | 25 | 22 | 44 |
| Intelligence | 44 | 48 | 52 | 74 | 69 | 28 | 59 | 37 | 58 | 37 | 51 | 37 | 27 | 67 | 67 |
| Sex | F | F | F | F | F | F | M | M | M | M | M | M | M | F | M |
| Frequency of scale position usage | | | | | | | | | | | | | | | |
| 1 and 7 | 17 | 29 | 30 | 22 | 31 | 34 | 29 | 13 | 8 | 8 | 9 | 11 | 33 | 4 | 26 |
| 4 | 24 | 20 | 11 | 24 | 9 | 15 | 7 | 19 | 0 | 31 | 7 | 28 | 16 | 27 | 20 |
| 2, 3, 5 and 6 | 29 | 21 | 39 | 24 | 30 | 21 | 34 | 38 | 62 | 31 | 54 | 31 | 21 | 39 | 24 |

# APPENDIX III

## RESIDUAL GROUP

| Subject No. | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 8 | 9 | 16 | 36 | 39 | 40 | 44 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Age | 44 | 23 | 21 | 23 | 33 | 28 | 40 | 39 | 15 | 27 | 37 | 29 | 39 | 21 |
| Intelligence | 54 | 63 | 42 | 34 | 48 | 26 | 56 | 60 | 32 | 72 | 60 | 46 | 50 | 31 |
| Sex | F | F | F | F | M | M | F | F | F | M | M | M | F | M |
| Frequency of scale position usage | | | | | | | | | | | | | | |
| 1 and 7 | 7 | 12 | 22 | 14 | 25 | 22 | 26 | 27 | 29 | 34 | 30 | 30 | 21 | 27 |
| 4 | 52 | 40 | 40 | 43 | 44 | 48 | 44 | 25 | 32 | 24 | 23 | 25 | 33 | 24 |
| 2, 3, 5 and 6 | 11 | 18 | 8 | 13 | 1 | 0 | 0 | 18 | 9 | 12 | 17 | 15 | 16 | 19 |

# APPENDIX IV

## Frequency of judging row 'better' than column for the Non-extreme group

| Concept No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | Row Sum |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Abortion | . | 0 | 5 | 1 | 11 | 2 | 0 | 1 | 0 | 5 | 0 | 11 | 1 | 37 |
| 2. Bath | 30 | . | 30 | 27 | 30 | 29 | 10 | 29 | 22 | 30 | 3 | 29 | 27 | 296 |
| 3. Danger | 25 | 0 | . | 3 | 24 | 9 | 1 | 2 | 1 | 11 | 0 | 29 | 2 | 107 |
| 4. Dark | 29 | 3 | 27 | . | 29 | 24 | 3 | 22 | 1 | 27 | 0 | 30 | 13 | 208 |
| 5. Devil | 19 | 0 | 6 | 1 | . | 3 | 1 | 2 | 1 | 5 | 0 | 15 | 0 | 53 |
| 6. Dirt | 28 | 1 | 21 | 6 | 27 | . | 4 | 9 | 3 | 19 | 0 | 27 | 4 | 149 |
| 7. Father | 30 | 20 | 29 | 27 | 29 | 26 | . | 29 | 23 | 28 | 5 | 30 | 27 | 303 |
| 8. Flea | 29 | 1 | 28 | 8 | 28 | 21 | 1 | . | 0 | 27 | 0 | 29 | 8 | 180 |
| 9. Lemon | 30 | 8 | 29 | 29 | 29 | 27 | 7 | 30 | . | 29 | 2 | 30 | 25 | 275 |
| 10. Pain | 25 | 0 | 19 | 3 | 25 | 11 | 2 | 3 | 1 | . | 0 | 28 | 2 | 119 |
| 11. Peace | 30 | 27 | 30 | 30 | 30 | 30 | 25 | 30 | 28 | 30 | . | 29 | 30 | 349 |
| 12. Torture | 19 | 1 | 1 | 0 | 15 | 3 | 0 | 1 | 0 | 2 | 1 | . | 0 | 43 |
| 13. Winter | 29 | 3 | 28 | 17 | 30 | 26 | 3 | 22 | 5 | 28 | 0 | 30 | . | 221 |

# APPENDIX IV

## Frequency of judging row 'worse' than column for the Non-extreme group

| Concept No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 Abortion | . | 29 | 26 | 30 | 14 | 26 | 30 | 29 | 30 | 27 | 30 | 17 | 30 |
| 2 Bath | 1 | . | 0 | 1 | 0 | 0 | 22 | 2 | 7 | 2 | 30 | 0 | 4 |
| 3 Danger | 4 | 30 | . | 28 | 8 | 26 | 30 | 25 | 30 | 23 | 30 | 4 | 28 |
| 4 Dark | 0 | 29 | 2 | . | 3 | 5 | 24 | 10 | 23 | 3 | 30 | 0 | 15 |
| 5 Devil | 16 | 30 | 22 | 27 | . | 24 | 30 | 30 | 30 | 23 | 30 | 12 | 28 |
| 6 Dirt | 4 | 30 | 4 | 25 | 6 | . | 29 | 16 | 29 | 10 | 30 | 0 | 27 |
| 7 Father | 0 | 8 | 0 | 6 | 0 | 1 | . | 0 | 11 | 1 | 25 | 0 | 6 |
| 8 Flea | 1 | 28 | 5 | 20 | 0 | 14 | 30 | . | 30 | 6 | 30 | 1 | 19 |
| 9 Lemon | 0 | 23 | 0 | 7 | 0 | 1 | 19 | 0 | . | 0 | 29 | 1 | 2 |
| 10 Pain | 3 | 28 | 7 | 27 | 7 | 20 | 29 | 24 | 30 | . | 30 | 2 | 29 |
| 11 Peace | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 1 | 0 | . | 0 | 0 |
| 12 Torture | 13 | 30 | 26 | 30 | 18 | 30 | 30 | 29 | 29 | 28 | 30 | . | 30 |
| 13 Winter | 0 | 26 | 2 | 15 | 2 | 3 | 24 | 11 | 28 | 1 | 30 | 0 | . |
| Column sum | 42 | 291 | 94 | 216 | 58 | 150 | 302 | 176 | 278 | 124 | 354 | 37 | 218 |

# APPENDIX IV

## Frequency of judging row 'better' than column for the Extreme group

| Concept No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | Row Sum |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Abortion | . | 2 | 10 | 3 | 15 | 4 | 3 | 2 | 3 | 11 | 2 | 14 | 4 | 73 |
| 2. Bath | 28 | . | 26 | 23 | 28 | 27 | 8 | 28 | 29 | 26 | 10 | 23 | 21 | 277 |
| 3. Danger | 20 | 4 | . | 7 | 24 | 10 | 2 | 10 | 4 | 16 | 3 | 26 | 10 | 137 |
| 4. Dark | 27 | 7 | 23 | . | 26 | 25 | 7 | 21 | 10 | 22 | 5 | 25 | 14 | 212 |
| 5. Devil | 15 | 2 | 6 | 4 | . | 5 | 2 | 6 | 3 | 5 | 2 | 12 | 3 | 65 |
| 6. Dirt | 26 | 3 | 20 | 5 | 25 | . | 2 | 14 | 3 | 19 | 2 | 23 | 6 | 147 |
| 7. Father | 27 | 22 | 28 | 23 | 28 | 28 | . | 27 | 28 | 26 | 13 | 28 | 24 | 302 |
| 8. Flea | 28 | 2 | 20 | 9 | 24 | 16 | 3 | . | 3 | 20 | 2 | 23 | 11 | 161 |
| 9. Lemon | 27 | 1 | 25 | 20 | 27 | 28 | 2 | 27 | . | 23 | 5 | 22 | 13 | 220 |
| 10. Pain | 19 | 4 | 14 | 8 | 25 | 11 | 4 | 10 | 7 | . | 2 | 19 | 6 | 129 |
| 11. Peace | 28 | 20 | 27 | 25 | 28 | 28 | 17 | 28 | 25 | 28 | . | 28 | 24 | 306 |
| 12. Torture | 16 | 7 | 4 | 5 | 18 | 7 | 2 | 7 | 8 | 11 | 2 | . | 6 | 93 |
| 13. Winter | 26 | 9 | 20 | 16 | 27 | 24 | 6 | 19 | 17 | 24 | 6 | 24 | . | 218 |

# APPENDIX IV

## Frequency of judging row 'worse' than column for the Extreme group

| Concept No. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 Abortion | . | 29 | 22 | 26 | 17 | 26 | 30 | 26 | 26 | 23 | 27 | 18 | 28 |
| 2 Bath | 1 | . | 2 | 3 | 3 | 4 | 23 | 1 | 10 | 2 | 22 | 2 | 4 |
| 3 Danger | 8 | 28 | . | 24 | 9 | 21 | 28 | 22 | 24 | 17 | 26 | 6 | 26 |
| 4 Dark | 4 | 27 | 6 | . | 3 | 7 | 25 | 8 | 22 | 10 | 27 | 4 | 15 |
| 5 Devil | 13 | 27 | 21 | 27 | . | 21 | 29 | 23 | 28 | 23 | 28 | 15 | 24 |
| 6 Dirt | 4 | 26 | 9 | 23 | 9 | . | 27 | 15 | 26 | 11 | 28 | 5 | 20 |
| 7 Father | 0 | 7 | 2 | 5 | 1 | 3 | . | 4 | 6 | 2 | 13 | 2 | 2 |
| 8 Flea | 4 | 29 | 8 | 22 | 7 | 15 | 26 | . | 27 | 10 | 29 | 5 | 16 |
| 9 Lemon | 4 | 20 | 6 | 8 | 2 | 4 | 24 | 3 | . | 4 | 27 | 5 | 4 |
| 10 Pain | 7 | 28 | 13 | 20 | 7 | 19 | 28 | 20 | 26 | . | 28 | 9 | 23 |
| 11 Peace | 3 | 8 | 3 | 3 | 2 | 2 | 17 | 1 | 3 | 2 | . | 2 | 3 |
| 12 Torture | 12 | 28 | 24 | 26 | 15 | 25 | 28 | 25 | 25 | 21 | 28 | . | 26 |
| 13 Winter | 2 | 26 | 4 | 15 | 6 | 10 | 28 | 14 | 26 | 7 | 27 | 4 | . |
| Column sum | 62 | 283 | 120 | 202 | 81 | 157 | 313 | 162 | 249 | 132 | 311 | 77 | 191 |

# APPENDIX V

## Frequency of 'better' judgements for individual subjects in the Non-extreme group

| Subject No. | 7 | 37 | 33 | 31 | 32 | 34 | 35 | 38 | 42 | 10 | 11 | 13 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Concept | | | | | | | | | | | | | | | |
| PEACE | 21 | 23 | 24 | 22 | 23 | 24 | 21 | 24 | 24 | 24 | 23 | 24 | 24 | 24 | 24 |
| FATHER | 14 | 18 | 22 | 24 | 23 | 19 | 24 | 16 | 22 | 22 | 21 | 14 | 20 | 22 | 22 |
| BATH | 18 | 21 | 20 | 18 | 19 | 20 | 21 | 20 | 20 | 20 | 20 | 21 | 21 | 20 | 17 |
| LEMON | 17 | 21 | 16 | 19 | 18 | 20 | 17 | 19 | 18 | 18 | 17 | 18 | 19 | 18 | 20 |
| WINTER | 15 | 17 | 13 | 13 | 12 | 14 | 15 | 13 | 12 | 16 | 18 | 17 | 14 | 16 | 16 |
| DARK | 9 | 13 | 16 | 17 | 13 | 12 | 15 | 12 | 16 | 14 | 15 | 18 | 14 | 11 | 13 |
| FLEA | 8 | 6 | 11 | 12 | 13 | 17 | 12 | 16 | 13 | 12 | 12 | 13 | 9 | 14 | 12 |
| DIRT | 16 | 12 | 13 | 8 | 4 | 7 | 6 | 13 | 11 | 10 | 10 | 10 | 14 | 7 | 8 |
| PAIN | 11 | 10 | 8 | 11 | 7 | 4 | 6 | 10 | 5 | 8 | 8 | 4 | 9 | 9 | 9 |
| DANGER | 5 | 8 | 7 | 5 | 15 | 8 | 10 | 6 | 7 | 6 | 6 | 2 | 6 | 9 | 7 |
| DEVIL | 8 | 5 | 0 | 5 | 5 | 6 | 2 | 0 | 0 | 0 | 3 | 7 | 4 | 2 | 6 |
| TORTURE | 6 | 2 | 4 | 0 | 4 | 0 | 7 | 2 | 5 | 4 | 3 | 0 | 2 | 4 | 0 |
| ABORTION | 8 | 0 | 2 | 2 | 0 | 5 | 0 | 5 | 3 | 2 | 0 | 8 | 0 | 0 | 2 |

(This data is from summing the rows on the Paired Comparisons matrices for individual subjects)

# APPENDIX V

## Frequency of 'worse' judgements for individual subjects in the Non-extreme group

| Subject No. | 7 | 37 | 33 | 31 | 32 | 34 | 35 | 28 | 42 | 10 | 11 | 13 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Concept | | | | | | | | | | | | | | | |
| PEACE | 23 | 24 | 24 | 22 | 23 | 24 | 24 | 24 | 23 | 24 | 24 | 24 | 23 | 24 | 24 |
| FATHER | 20 | 22 | 22 | 24 | 23 | 21 | 22 | 16 | 23 | 20 | 18 | 15 | 12 | 22 | 22 |
| BATH | 21 | 16 | 19 | 18 | 20 | 17 | 20 | 20 | 19 | 20 | 19 | 21 | 21 | 20 | 20 |
| LEMON | 19 | 17 | 17 | 18 | 18 | 21 | 16 | 21 | 17 | 18 | 22 | 17 | 21 | 18 | 18 |
| WINTER | 13 | 15 | 14 | 16 | 15 | 10 | 15 | 15 | 14 | 12 | 15 | 16 | 17 | 16 | 15 |
| DARK | 7 | 12 | 14 | 16 | 15 | 13 | 17 | 13 | 17 | 18 | 11 | 21 | 15 | 14 | 13 |
| FLEA | 11 | 8 | 7 | 10 | 7 | 17 | 12 | 16 | 12 | 14 | 13 | 11 | 14 | 10 | 14 |
| DIRT | 12 | 11 | 14 | 11 | 8 | 10 | 7 | 9 | 11 | 10 | 10 | 6 | 11 | 12 | 8 |
| PAIN | 11 | 12 | 9 | 9 | 11 | 6 | 7 | 8 | 4 | 8 | 9 | 6 | 8 | 8 | 8 |
| DANGER | 11 | 5 | 7 | 6 | 10 | 7 | 8 | 8 | 7 | 6 | 6 | 4 | 4 | 6 | 6 |
| DEVIL | 6 | 11 | 0 | 2 | 3 | 4 | 3 | 0 | 0 | 1 | 7 | 8 | 8 | 2 | 3 |
| ABORTION | 7 | 1 | 4 | 3 | 0 | 6 | 1 | 3 | 3 | 1 | 0 | 6 | 0 | 2 | 5 |
| TORTURE | 2 | 2 | 5 | 1 | 3 | 0 | 4 | 3 | 6 | 4 | 2 | 1 | 2 | 2 | 0 |

(This data is from summing the columns on the Paired Comparisons matrices for individual subjects)