MATHEMATICAL MODELS OF THE RELIABILITY OF THE SEMANTIC DIFFERENTIAL

Thesis for the Degree of M. A.

MICHIGAN STATE UNIVERSITY
THOMAS SCOTT NICOL

1972

ABSTRACT

MATHEMATICAL MODELS OF THE
RELIABILITY OF THE SEMANTIC DIFFERENTIAL

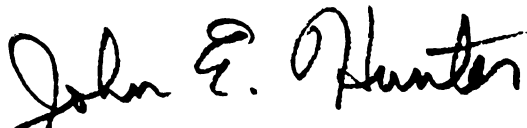
By

Thomas Scott Nicol

This paper is devoted to the internal consistency of the semantic differential. The initial hypothesis was that the usual method of presenting all the adjectives for a single concept at the same time might lead to reliability data with "correlated errors." In grappling with the data a series of four formal mathematical models were created and tested against the data. The model which fit best assumed (1) that subjects draw on only a sample of their belief system in responding to the adjectives, (2) that the usual semantic differential format causes subjects to create idiosyncratic definitions which "overdifferentiate" the evaluative adjectives, and (3) that some of the subjects draw only one cognitive sample to make all the responses while others draw a new sample for each response. The data collected were interpreted as showing a definite "correlated errors" component to the usual semantic differential format. The discussion noted that for high

population homogeneity on a concept, this could result in a very substantial "spuriousness" in reliability estimates using any measure of internal consistency.

Approved May, 1972



John E. Hunter, Chairman

MATHEMATICAL MODELS OF THE RELIABILITY OF THE SEMANTIC DIFFERENTIAL

By

Thomas Scott Nicol

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF ARTS

Department of Psychology

1972

ACKNOWLEDGMENTS

I would like to express my appreciation to the following people for the aid and support which they contributed to this project; John Hunter for his unwavering and invaluable assistance from inception to completion, Ralph Levine for his critical reading of the first draft and finally my wife Julie, without whose patience and support this project would have been impossible.

TABLE OF CONTENTS

Page

APPENDIX

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The use of bipolar adjective scales as a psychological measurement technique developed out of a series of studies on color-music synthesis by Karwoski, Eckerson, Odbert, Osgood and others during the late 1930's and early 1940's. In 1946, Stagner and Osgood (1) adapted the use of bipolar scales to the study of social stereotypes. Between 1946 and 1957, Osgood and others performed a number of factor analytic studies using a set of bipolar adjectives scales which they referred to as a semantic differential. At this time, Osgood (2) proposed that the semantic differential technique might be used as a means for measuring the meaning of concepts.

These studies by Osgood and others culminated in 1957 with the publication of The Measurement of Meaning by Osgood, Suci and Tannenbaum (3). The dominant result of these factor analytic studies was the consistent finding of three major dimensions in the semantic space: evaluative, activity and potency. The authors proposed that the evaluative dimension of the semantic space might be used as a method for measuring attitudes. Since that time many researchers have adopted this suggestion.

Semantic Differential Formats

There are a number of formats which the semantic differential may take. By far the most popular is what will be called the fixed concept form. In this format, the subject is asked to respond to all of the adjective scales for one concept before proceeding to the next concept. When using this format to measure attitudes, Osgood et. al. (3) recommends the use of adjective scales from nonevaluative dimensions along with those from the evaluative in order to obscure the purpose of the instrument. No studies were found which compared the results of using the evaluative scales alone to the suggested mixing of scale dimensions.

Osgood et. al. (3) reports a study of one other format of the semantic differential. For this format, the adjective scale-concept items are presented in a random order. In the study by Kerrick, 10 adjectives and eight concepts were presented to the same group of subjects twice; once using the fixed concept form and once using the random form. Comparing the two forms, Kerrick found that only three of the 80 scale concept means were significantly different. The authors concluded that it makes no difference which format is used.

Wells and Smith (4) were also concerned with the question of format. They compared the fixed concept format with a fixed adjective format. In the fixed adjective format, responses to all the concepts for a single

adjective scale are made before proceeding to the next adjective scale. They found greater differences between concepts for the fixed adjective format than for the fixed concept format.

Reliability Studies

As a portion of the first factor analytic study reported by Osgood et. al. (3) 40 adjective scale-concept items were administered to the subjects twice with no time interval between administrations. The correlation between the two sets of responses pooled over the 40 items was .85. A study by Tannenbaum is also reported in which six concepts and six adjective scales, 36 items, were administered twice to a group of subjects. Correlations were obtained for each subject between his six concept sum scores for the two tests. These correlations ranged from .87 to .93 with a mean of .91.

Various other studies have been concerned, either directly or indirectly, with the topic of test-retest reliability. Jenkins, Russell and Suci (5) correlated mean adjective scale values for 20 scales and 20 concepts. They report a test-retest correlation of .97 for a retest interval of four weeks. Using 20 concepts, 20 adjective scales and an immediate retest, Miran (6) reports a test-retest correlation of .99 between scale means for each concept.

In a study concerned with the test-retest reliability

of children's ratings, Divesto and Dick (7) found that the correlations between each response increased as the age of the child increased. For a test-retest interval of three to four weeks, the average correlation ranged from .26 for children in the third grade to .55 for those in the seventh grade. The immediate test-retest average correlations were .56 and .67 for the same two grades.

The only report found using an internal consistency measure to estimate reliability was done by Block (8). Block reports a split-half reliability of .70 using two concepts and eighty adjective scale.

The present study. The present study was motivated by two desires: first, to make a more comprehensive study of the internal consistency of the semantic differential than is in the current literature; and second, to test some hypotheses concerning the relation between format and reliability. However, after the data was analyzed, the results were so different from what was anticipated that the study took a different turn. That is, in searching for an explanation of the results, an entire class of relatively formal models of the response processes underlying the semantic differential were generaged. These models will be presented and tested below.

MODELS

Introduction

This paper will present a number of models of the response producing process generated by semantic differential items. Before proceeding with the development of these models, the nature of the data with which they will deal needs to be clarified.

The models will be concerned with responses to bipolar adjective scales for a single concept. Furthermore, the adjective scales considered will be assumed to represent a single dimension of the semantic space and to be mutually parallel. That is, while the scales actually consist of different adjectives and while the subjects might feel that they have distinctively different meanings, it is assumed that for a given subject and concept, the scales would all arouse identical response processes. Two adjectives from such a set of scales will be called functionally synonymous. In the data, the adjectives were all chosen to be synonomous with good-bad, i.e. Osgood's evaluative dimension.

If the adjectives are in fact functional synonyms, then the basic data should satisfy the assumptions of classical reliability. Thus for a given concept each

adjective should have the same mean and variance, and the intercorrelations between any pair of adjectives should be the same as the correlation between any other pair. If the item means and variances in the data are approximately equal (as in the present data) and if the interadjective correlations within a given concept are all about the same size, i.e. the interadjective correlation matrix is flat, then the observed responses X_i are all related to the same common factor F by the equation

$$X_i = F + u_i$$

where u_i is the "unique" part of X. Furthermore $r_{Fu_i} = r_{u_i u_j} = 0$ for all the unique factors and all the unique factors have the same variance. If there were infinitely many functional synonyms the observed scores could be averaged to form

$$\overline{X} = F + \overline{e} \rightarrow F + 0 = F$$

This last theorem is particularly important since it constitutes at least an "in principle" identification of the hypothetical variables F and u_i.

The theorems in the previous paragraph are all exactly analagous to the familiar theorems based on the assumption that the observed variables X_i are based on the same true score T and satisfy the classical error equation

$$X_i = T + E_i$$

where the E_i's are the Spearman (9) errors that are correlated with nothing on Earth save themselves. Thus it is very tempting to assume that T = F and $E_i = u_i$. If this were true, then the reliability of each adjective (for the given concept) would be the interadjective correlation r_{xx} and the reliability of the evaluative score would be calculated by coefficient alpha, Cronbach (10), and would reduce for this data to the classical Spearman-Brown formula. Furthermore this internal consistency coefficient would be equal to the "immediate" test-retest coefficient and would be suitable for use in correcting interscale correlations for attenuation.

Unfortunately there are several problems in making the identification of F with T. Mathematically the main problem lies in the statement errors are uncorrelated with any other independently defined variable. The basis of all attenuation formulas is the assumption that $\sigma_{xy} = \sigma_{Fy}$. This in turn is true only if $\sigma_{uy} = 0$, i.e. the "errors are uncorrelated..." assumption. Actually this is a plausible assumption in the case of the semantic differential. It seems very unlikely that the processes which produce the difference between two successive semantic differential responses would play a significant role in the scheme of things. Thus mathematically F can indeed play the role of a "true score." Correction for attenuation using coefficient alpha would convert correlations from r_{xy} to r_{Fy}.

However psychologically F might not be a satisfactory "true score." Consider the following hypothetical example. Suppose that the subject's mood varies randomly from day to day. Assume further that the subject is in a "good" mood when his mood for that day is higher than his average. Finally assume that when a subject is in a good mood he gives a more positive evaluation to any object than if he were neutral and that he gives uniformly lower evaluations when he is in a bad mood. That is, the common factor F will vary with the subject's mood. A psychologist would probably want to define the "true score" to be the subject's response when in a neutral mood. Denote this response by T. If mood m is measured by the deviation from the subject's own average, then

$$F = T + m$$

If we equate

$$X_i = T + e_i = F + u_i$$

then we have

$$e_i = m + u_i$$

Now

$$r_{e_i e_j} = \frac{\sigma_m^2}{\sigma_X^2} \neq 0$$

so we have a case of measurement with "correlated errors."

If T is regarded as "true score," then coefficient alpha is not the proper reliability coefficient since

$$\text{alpha} = r_{XF}^2 > r_{XT}^2$$

Furthermore "correction for attenuation" yields spuriously low correlations:

$$|r_{yF}| < |r_{yT}|$$

The hypothesis which generated this study was that the errors in the usual fixed concept format occur so closely in time that they might be very highly correlated.

Three special formats

The most popular class of formats is the one in which the subject is asked to respond to all of the adjectives for one concept before proceeding to the next concept. That is, all responses to one concept are obtained before another concept is considered. These semantic differential formats will be called fixed concept formats; one concept is "fixed" and all the adjective scales are presented, then another concept is "fixed" and the scales presented again, etc.

This class of formats will be divided into two subclasses. One type of fixed concept format which will be considered is one in which only functionally synonymous scales are used. For example, these scales might represent Osgood's evaluative dimension of the semantic space. In this study, this format will be called the fixed concept, evaluative only format.

The other type of fixed concept format which will be considered is one in which the same evaluative scales

are intermixed with scales from other semantic dimensions. The other scales will not be functionally synonymous with the evaluative set. That is, the additional scales are specifically designed to ask the subject to rate the particular concept on a nonevaluative basis. For example, while a subject may feel that labor unions are bad, (evaluative) he might also feel that they are growing (non-evaluative) or he might feel that they are important (non-evaluative), etc. The non-evaluative scales would not be expected to correlate highly with the evaluative scales. This second fixed concept format will be called the fixed concept, multiple dimension format.

Finally, the third format which the models will consider is one which is the transpose of the fixed concept format. In these formats, all of the concepts are rated on a single scale before they are rated on another scale. That is, one scale is presented and all of the concepts are rated on this scale. Then another scale is presented and all of the concepts are rated on this scale. Thus, the adjective scale is held "fixed" while the concepts are varied and the name given is fixed adjective format.

As was indicated earlier, the models which will follow were developed to describe responses to semantic differential items. These models will make various assumptions about the response formation process and

how this process interacts with each of the three types of formats described above. These models will all start from the classical basis that each response to an item is composed of true score plus "error." For the most part, the interaction between the response forming process and the three formats will have its greatest impact upon the composition and nature of the "error" term. Thus, within a given model, the assumptions about response formation will have different effects upon each of the three formats.

Pure Response Error

Suppose that the subject has just been presented with a concept and an adjective pair. The first thing that happens is that by some process the subject has an internal response to the item. Let this response or "feeling" be denoted F. The subject is then required to translate this evaluation into one of a given set of limited response categories.

Suppose that after looking over the response categories, the subject feels that none of them is exactly appropriate. For example, suppose that the subject would like to translate his feelings into a response which falls between two categories. In a sense, the subject's "true" response falls between the two response categories. Since the subject is required to choose one of the available responses, he is forced to

introduce error into his actual response. His actual response will not be the same as his true response would have been. If X represents the subjects actual response, F his true response and E the difference between his true response and actual response then $X = F + E$.

From the subject's point of view, since neither of the response categories accurately represents his true feelings, they are both equally good (or bad) responses. The subject must choose one of these responses and since he has no reasonable basis on which to make the choice, he makes a random decision. Because the subject's choice is random, the error term E, is also random.

If the responses from a number of subjects to a particular concept-scale item are considered, the error in each of their responses due to the translation procedure is random. One consequence of this is that the errors are unrelated to the subjects "true" feelings. This can be clarified if we consider a simple example of responses from two subject's within the response categories 1,2,3,4,5,6,7. Subject one feels that his "true" response falls between 2 and 3. He therefore makes a random choice between the two. If he chooses the 2, his actual response is too low, the error is negative. If he chooses the three, his response is too high, the error is positive. Subject two feels that

his "true" response falls between 6 and 7 so he too must make a choice. If he chooses the 6, his response is too low or the error is negative. If he chooses the 7, his response is too high or the error is positive. From this simple example, it should be clear that the sign of the error term, E, is unrelated to the subject's "true" response, F. This relation can be expressed by $r_{EF} = 0$.

What about the error in responses to two different scale items for the same concept. Again, the error in either one is completely random. Therefore, if responses from a group of subjects are considered, the errors should be uncorrelated with each other. That is, if E_i represents the errors in responses to one item and E_j the errors in the other, then $r_{E_iE_j} = 0$.

Model I: Direct Response Elicitation

The first model assumes that the subject's evaluative response is a true affective response, i.e. an emotional response that is elicited by the adjective-concept item without cognitive deliberation. If T is the subject's pure emotional response to the item, then one version of the affective assumption is simply

$$F = T$$

The observed response is then given by

$$X = F + E = T + E$$

and model I is shown to be mathematically equivalent to classical error theory.

To make this point more precisely, let X_1 and X_2 be the responses to two functionally synonomous adjective scales. Then $X_1 = T + E_1$ and $X_2 = T + E_2$. The expected value of functionally synonymous scales are the same for all the scales on a given concept. This equality also holds for the variances of these scales. That is, $E(X_1) = E(X_2)$ and $\sigma^2_{x_1} = \sigma^2_{x_2}$.

Since the correlation between T and E is zero, the variance of these two concept-scale items can be expressed as follows:

$$\sigma^2_{x_1} = \sigma^2_T + \sigma^2_{E_1}$$

and

$$\sigma^2_{x_2} = \sigma^2_T + \sigma^2_{E_2}$$

Also, the covariance of the two functionally synonymous scales is:

$$\sigma_{x_1x_2} = \sigma_{TT} + \sigma_{TE_2} + \sigma_{E_1T} + \sigma_{E_1E_2}$$

Again, because the error terms are uncorrelated with either T or with each other, this reduces to

$$\sigma_{X_1X_2} = \sigma_{TT} = \sigma^2_T$$

The correlation between the two functionally synonymous scales is given by the usual formula

$$r_{X_1X_2} = \frac{\sigma^2_T}{\sigma^2_x}$$

Another familiar formula emerges from this model

$$r_{XT} = \frac{\sigma_T}{\sigma_X}$$

or $$r_{XX} = r_{XT}^2$$

and coefficient alpha would be appropriate for this model.

For reasons that will be given later, one last statistic will be noted and referred to as the unique variance.

$$\sigma_X^2 (1-r_{X_1X_2}) = \sigma_X^2 - \sigma_{X_1X_2} = \sigma_E^2$$

The process which produces the random error, E, is independent of the content of the concept. Thus, σ_E^2 should be the same for all concepts.

Model II: An Information Processing Model

The previous model assumes that a subject's feelings are directly given or elicited. However this may not be true. Suppose instead that a subject constructs his response on the basis of his beliefs and ideas about the object. Then before a subject can make a response, he must first evaluate the information which is currently present in his cognitive domain. In this model, it will be assumed that the amount of information currently present in the subject's cognitive domain relevant to any given concept is too large and complex for him to base his response on all this information. Instead, he

takes a random sample of the information in his cognitive domain and uses an evaluation of this sample as the basis for his response. Since the evaluative content of any given sample may be either higher or lower than the evaluative content of the entire domain, this sampling process introduces a new element into the subject's feelings. Ideally the psychologist would like to know the subject's evaluative response to the entire domain. So this hypothetical response would be the "true score" and will be denoted T. Let the deviation of the evaluative content of the sample from the evaluative content of the domain be denoted e. Then the subject's feelings are given by

$$F = T + e$$

and the observed response will be given by

$$X = T + e + E$$

where E is the same response error as in model I. From the psychologist's point of view, the e + E is a sort of differentiated stepchild of the "error" in classical reliability theory.

How do the new variables correlate with one another? Since T is a constant for a given subject while e varies randomly, T and e are uncorrelated within a given subject. But if e is randomly related to T within subjects it will be randomly related across subjects, so $r_{Te} = 0$. Finally since E is the error generated by the subject's attempt to translate his

feelings into a response category, E will be independent of both T and e, i.e. $r_{eE} = r_{TE} = 0$. Thus we have a formula for the variance

$$\sigma_X^2 = \sigma_{T+e}^2 + \sigma_E^2 = \sigma_T^2 + \sigma_e^2 + \sigma_E^2$$

However there is a decision to be made before the correlation between two responses can be obtained: Is the subject's new response based on a new sample from the cognitive domain or on the old sample? Suppose he resamples the appropriate portion of his cognitive domain. The sample of information which a subject takes from his cognitive domain is a random sample and therefore depends only on the current content of the domain and not on how it came to be that way. In particular, the sample taken for a second response will be independent of that taken for the first response. Thus if the two responses are

$$X_1 = T + e_1 + E_1$$

and

$$X_2 = T + e_2 + E_2,$$

then $r_{e_1e_2} = 0$. The covariance is then given by

$$\sigma_{x_1x_2} = \sigma_{TT} = \sigma_T^2$$

Hence

$$r_{X_1X_2} = \frac{\sigma_T^2}{\sigma_X^2} = r_{XX}$$

and

$$r_{XX} = r_{XT}^2$$

which are again the formulas for classical reliability. And for future reference, the unique variance is

$$\sigma_X^2 (1-r_{XX}) = \sigma_X^2 - \sigma_{X_1X_2} = \sigma_e^2 + \sigma_E^2$$

However suppose the subject doesn't base his second response on a new sample from the domain. For example, in the fixed concept format, the subject answers a sequence of questions about the same object. He could use the same sample for all the answers. If so, the second deviation e_2 would exactly equal the first e_1. That is, the two observed responses would be

$$X_1 = T + e + E_1$$
$$X_2 = T + e + E_2$$

The covariance and correlation would then be

$$\sigma_{X_1X_2} = \sigma_{T+e,T+e} = \sigma_T^2 + \sigma_e^2$$

$$r_{X_1X_2} = \frac{\sigma_T^2 + \sigma_e^2}{\sigma_X^2} = r_{XX}$$

which is not the classical formula. Furthermore

$$r_{XT} = \frac{\sigma_T}{\sigma_X}$$

or

$$r_{XT}^2 = \frac{\sigma_T^2}{\sigma_X^2} \neq r_{XX}$$

Thus, model IIb is a correlated errors model and coefficient alpha would not be an appropriate estimate of the reliability.

Finally, for future reference, the unique variance is

$$\sigma_X^2 (1-r_{XX}) = \sigma_X^2 - \sigma_{X_1X_2} = \sigma_E^2$$

For the fixed adjective format, the subject is forced to take a new sample for each adjective. However for the fixed concept format, it is conceivable that one sample would be used for all responses. This suggests two models. In model IIa, the subject always resamples even for fixed concept formats. In model IIb, the subject resamples for the fixed adjective format but uses only one sample for the fixed concept formats.

There is a third possible model. The subject might be assumed to resample for the fixed adjective and fixed concept, multiple dimension formats, but not resample for the evaluative only format. But a careful examination suggests that this is very implausible. If the subject resamples in the multidimensionale case, it is because one of two processes "erased" the sample: the conversion from feeling to response category or the reaction to the new adjective-dimensions. But both of these processes are also at work in the fixed concept, evaluative only format. Although the adjectives may be <u>functional</u> synonyms, they are different words and the subject will react to them as different questions.

One last point deserves mention. The value of σ_E^2 is the same for all concepts, the value of σ_e^2 is not.

The greater the cognitive complexity of the domain for a given concept, the greater the value of σ_e^2 for that concept.

Model III: Forced Variance Model

The development of this model begins with a consideration of the nature of the fixed concept, evaluative only format. For this format only the functionally synonymous adjectives are rated. Thus, this format is intentionally designed to repeatedly ask the same question over and over. This means that a subject's responses to this set of questions vary only to the extent that the errors produce different responses. Thus from the subject's viewpoint, all of his responses for a given concept are highly similar. If this lack of variance makes the subject feel uncomfortable or vaguely guilty about his performance, then the subject might well regenerate responses until they show more variation. That is, each "natural" response would be arbitrarily changed by some amount. Since these changes are unrelated to the subject's feelings about the concept, they are "errors."

Let this forced error in a given response be represented by ϵ. Combining this forced error term with the error terms E and e, introduced in models I and II results in the representation of a single response by $X = T + e + E + \epsilon$.

Since the response change, ϵ, which subjects make are arbitrary, the error introduced by these changes is uncorrelated with the subjects' true feeling, i.e. $r_{T\epsilon} = 0$. Furthermore, because ϵ is abritrary, it is also uncorrelated with either E or e, $r_{E\epsilon} = r_{e\epsilon} = 0$. Using this information, the following expression for the variance of a response is produced.

$$\sigma_X^2 = \sigma_T^2 + \sigma_{e+\epsilon+E}^2 = \sigma_T^2 + \sigma_e^2 + \sigma_\epsilon^2 + \sigma_E^2$$

Model II, introduced the question of whether or not a subject resamples his cognitive domain for each response. The hypothesis of arbitrary response changes introduced here is compatable with either of these cognitive sampling assumptions. Thus, the assumption of forced variance in the evaluative only fixed concept format can be added to model IIa to produce model IIIa or to model IIb to produce IIIb.

Consider first model IIIa, which is based on model IIa, the resampling model. The representation of a subject's responses to two functional synonyms is

$$X_1 = T + e_1 + E_1 + \epsilon_1$$

$$X_2 = T + e_2 + E_2 + \epsilon_2$$

Since both ϵ_1 and ϵ_2 are arbitrary changes, $r_{\epsilon_1\epsilon_2} = 0$. Furthermore, since the forced response change is arbitrary and is unrelated to the cognitive samples, $r_{e_1\epsilon_2} = 0$ and $r_{e_1\epsilon_2} = 0$. Thus, the covariance is given by

$\sigma_{X_1X_2} = \sigma_{TT} = \sigma_T^2$. Therefore,

$$r_{X_1X_2} = \sigma_T^2/\sigma_X^2 = r_{XX}$$

and

$$r_{XX} = r_{XT}^2$$

which are the formulas for classical reliability. The following is again given for future reference:

$$\sigma_X^2 (1 - r_{X_1X_2}) = \sigma_X^2 - \sigma_{X_1X_2} = \sigma_\epsilon^2 + \sigma_\epsilon^2 + \sigma_E^2$$

Model IIIb can now be obtained from model IIb which assumes that both responses are based on the same cognitive sample. For model IIIb the observed responses to the functional synonyms would be given by

$$X_1 = T + e + \epsilon_1 + E_1$$

$$X_2 = T + e + \epsilon_2 + E_2$$

The covariance and correlation would then be:

$$\sigma_{X_1X_2} = \sigma_{T+e,T+e} = \sigma_T^2 + \sigma_e^2$$

and

$$r_{X_1X_2} = \frac{\sigma_T^2+\sigma_e^2}{\sigma_X^2} = r_{XX}$$

which are not the classical reliability formulas. Furthermore in this case

$$r_{XT}^2 = \frac{\sigma_T^2}{\sigma_X^2} \neq r_{XX}$$

Again, for purposes of future reference, the unique variance is

$$\sigma_X^2 (1-r_{X_1X_2}) = \sigma_X^2 - \sigma_{X_1X_2} = \sigma_\epsilon^2 + \sigma_E^2$$

Up to this point, the model has been developed by considering the fixed concept, evaluative only format. Thus either of the two sub models presented above applies directly to this format. What about responses from the fixed adjective format. Subject who respond to the fixed adjective format never see all of their responses to a given concept together. So, in the fixed adjective format, the subject never notices the similarity of responses to a given concept and thus feels no need to introduce an arbitrary change in these responses.

In the fixed concept, multiple dimension format, the subject not only responds to the set of functionally synonymous adjectives but also to a set of adjectives which are specifically designed to tap other dimensions. The responses to these adjectives would introduce considerable natural variance into a given subject's set of responses to a given concept. Thus, the subject would feel no need to introduce artificial variation.

There is one final complication to be considered. What is the relationship, if any, between σ_e^2 and σ_ϵ^2 ? For model IIIb, the subject uses only one sample. Thus from his point of view, the natural variance in his responses to any one concept would be approximately σ_E^2. Hence σ_ϵ^2 would be the same for all concepts for model IIIb. On the other hand, in model IIIa the subject draws a new sample for each response. Thus the natural variance that the subject sees should be $\sigma_e^2 + \sigma_E^2$.

Since σ_e^2 varies with the complexity of the domain, the variance that the subject sees will be different for different concepts. Since the amount of forced variance depends on the amount of observed variance, σ_ϵ^2 and σ_e^2 will be negatively correlated across concepts in model IIIa.

Model IV: Semantic Overdifferentiation

Consider again a subject who is in the process of responding to the fixed concept, single dimension format of the semantic differential. Again, suppose that the subject feels that he is not producing enough variance in his responses. The subject looks at the list of adjectives and notices that they all seem to be very similar in meaning. This apparent similarity of meaning bothers the subject because he feels that no one would expect him to reply to the same question over and over. Thus, he feels that he must not be making a fine enough discrimination. So he looks at the particular concept in question and invents idiosyncratic definitions of each adjective scale that apply to only the given concept. What he has done is taken a carefully prepared set of functional synonyms and created a temporary semantic differentiation among them.

When the subject draws his sample of cognitive elements, there is a subset of them which he would have used to generate his response had the adjective occurred

in isolation. This is the set of elements which are relevant to the evaluative dimension and would be the same for all adjectives. But, his semantic over-differentiation causes him to exclude some of these elements because they apply more "aptly" to some other adjective. Furthermore, when the subject is looking for elements which apply specifically to the idiosyncratic definition of the adjective he has invented, he may introduce previously excluded cognitive elements.

The mathematization of this model will parallel the process itself. As before the evaluative content of the entire domain is T. The content of a sample is T + e, where e is the deviation of the sample from the domain. Now T + e is the value that the subject would assign to this sample for any adjective had the adjectives occurred in isolation. However, if he modifies the sample by overdifferentiation he will obtain a new value which can be represented by T + e + f where f is now the deviation from the value assigned to the adjective in isolation. Since the new definitions are arbitrarily created, they will be independent of evaluative content when considered for different subjects. Thus across subjects the mean value of f is 0 and f is independent of T and e. Since the error produced by converting a feeling to a pencil mark is still present in this model, the full equation is

$$X = T + e + f + E$$

Since f is independent of the other components the variance

is given by

$$\sigma_X^2 = \sigma_T^2 + \sigma_{e+f+E}^2 = \sigma_T^2 + \sigma_e^2 + \sigma_f^2 + \sigma_E^2$$

In calculating the covariance of two responses, the critical question again is whether or not a new sample is drawn.

If a new sample is drawn then

$$X_1 = T + e_1 + f_1 + E_1$$

$$X_2 = T + e_2 + f_2 + E_2$$

and the covariance and correlations are given by

$$\sigma_{X_1X_2} = \sigma_{TT} = \sigma_T^2$$

$$r_{X_1X_2} = \frac{\sigma_T^2}{\sigma_X^2} = r_{XX}$$

In addition

$$r_{XT}^2 = \frac{\sigma_T^2}{\sigma_X^2} = r_{XX}$$

we have the familiar equations of classical reliability theory. The statistic for future reference is

$$\sigma_X^2 (1-r_{XX}) = \sigma_X^2 - \sigma_{X_1X_2} = \sigma_e^2 + \sigma_f^2 + \sigma_E^2$$

If the same sample is used for both responses, then

$$X_1 = T + e + f_1 + E_1$$

$$X_2 = T + e + f_2 + E_2$$

These equations are notable for the fact that they contain two independent components f_1 and f_2 even though one sample was drawn. The covariance is

$$\sigma_{X_1X_2} = \sigma^2_{T+e} = \sigma^2_T + \sigma^2_e$$

which again is the result for "correlated errors." The two correlations of interest are

$$r_{X_1X_2} = \frac{\sigma^2_T + \sigma^2_e}{\sigma^2_X} = r_{XX}$$

and

$$r^2_{XT} = \frac{\sigma^2_T}{\sigma^2_X} \neq r_{XX}$$

which means that coefficient alpha would not be appropriate. The statistic being computed for future interest, the unique variance, is

$$\sigma^2_X (1-r_{XX}) = \sigma^2_X - \sigma_{X_1X_2} = \sigma^2_f + \sigma^2_E$$

The present discussion has been predicated on the assumption of an evaluative only, fixed concept format. Before the number of models can be determined, the discussion must be extended to the other formats. How are the other two formats effected by the new response formation process introduced into this model? In the fixed adjective format, the subject has no opportunity to make a comparison of the adjectives. Each response to an evaluative adjective for a given concept is separated by responses to other concepts on the same adjective and by responses to all of the concepts on a non-synonymous adjective. Thus the subject will not be lead to create idiosyncratic definitions. That is,

responses to functionally synonymous adjectives in the fixed adjective format could be expressed as

$$X = T + e + E$$

which is, of course, the same expression as model II.

The effect of semantic over-differentiation on the fixed concept, multiple dimension is more complicated and depends upon the assumption about resampling. If the subject resamples his cognitive domain before each response, he probably won't notice that some of the adjectives are functionally synonymous. The intervening responses to adjectives from other semantic dimensions and the sampling procedure itself are probably enough to make the subject react to each adjective independently.

However, if the subject takes only a single cognitive sample and proceeds to base all of his responses on the elements in this sample, then he evaluates these elements in relationship to all of the adjectives. Thus he is more likely to notice that a particular sub-set of the adjectives, those from the evaluative dimension are all making use of the same elements. Thus, like the subject who responds to the fixed concept, single dimension format, this subject feels that he is not discriminating among the adjectives and proceeds to over-differentiate them.

Thus when all is said and done, there are two models: a model for subjects who draw a single sample and one for subjects who resample. Model IVa assumes

that subjects resample. For this model, the fixed adjective and fixed concept, multiple dimension formats are described by the equation of model II

$$X = T + e + E$$

while his response to the evaluative only fixed concept format is

$$X = T + e + f + E$$

If the subject makes multiple responses from the same sample, then his response to the fixed adjective format is

$$X = T + e + E$$

while his response to both fixed concept, evaluative only formats is

$$X = T + e + f + E$$

The last question to be discussed is the relationship between σ_f^2 and σ_e^2. For a single subject, the absolute value of f will be large only if two things are true. First, the subject must be able to generate distinct meanings for the synonyms for that concept. Second, the differences in content must yield differences in evaluation. Both of these conditions are strongly dependent on the cognitive complexity of the domain in question. Thus σ_f^2 and σ_e^2 should both be positively correlated with cognitive complexity and hence with each other.

DIFFERENTIATING THE MODELS

Differentiating the Models Using a Single Statistic

Four models of the response formation process have been presented for three forms of the semantic differential. All of these models have been mathematically similar to classical reliability theory in that each response is composed of true score plus "error." Each model broke the error component into finer categories. Table 1 presents the equations of two parallel responses developed for each of the models.

Using these response equations and the appropriate independence assumptions for each model, the variance of an item was expressed in terms of its components. These variances are presented in Table 2. The last column of this table displays the relationships among the item variances for each format predicted by the different models. For example, models I, IIa, and IIb all predict that a given item variance will be the same for all three formats. Actually, the rank order of the variances is the only observable feature of the table. Thus if the models are to be differentiated they must be differentiated on the basis of these rank orders.

In the next table, Table 3, the three predicted

Table 1.--Response equations developed from each of the models for the three formats.

	Formats		
	Fixed adjective	Fixed concept multiple dimension	Fixed concept evaluative only
I	$X_1 = T+E_1$	$X_1 = T+E_1$	$X_1 = T+E_1$
	$X_2 = T+E_2$	$X_2 = T+E_2$	$X_2 = T+E_2$
II Resampling	$X_1 = T+e_1+E_1$	$X_1 = T+e_1+E_1$	$X_1 = T+e_1+E_1$
	$X_2 = T+e_2+E_2$	$X_2 = T+e_2+E_2$	$X_2 = T+e_2+E_2$
II No Resampling	$X_1 = T+e_1+E_1$	$X_1 = T+e+E_1$	$X_1 = T+e+E_1$
	$X_2 = T+e_2+E_2$	$X_2 = T+e+E_2$	$X_2 = T+e+E_2$
III Resampling	$X_1 = T+e_1+E_1$	$X_1 = T+e_1+E_1$	$X_1 = T+e_1+\epsilon_1+E_1$
	$X_2 = T+e_2+E_2$	$X_2 = T+e_2+E_2$	$X_2 = T+e_2+\epsilon_2+E_2$

Table 1.--Continued.

	Formats		
	Fixed adjective	Fixed concept multiple dimension	Fixed concept evaluative only
III No Resampling	$X_1 = T+e_1+E_1$ $X_2 = T+e_2+E_2$	$X_1 = T+e+E_1$ $X_2 = T+e+E_2$	$X_1 = T+e+\varepsilon_1+E_1$ $X_2 = T+e+\varepsilon_2+E_2$
IV Resampling	$X_1 = T+e_1+E_1$ $X_2 = T+e_2+E_2$	$X_1 = T+e_1+E_1$ $X_2 = T+e_2+E_2$	$X_1 = T+e_1+f_1+E_1$ $X_2 = T+e_2+f_2+E_2$
IV No Resampling	$X_1 = T+e_1+E_1$ $X_2 = T+e_2+E_2$	$X_1 = T+e+f_1+E_1$ $X_2 = T+e+f_2+E_2$	$X_1 = T+e+f_1+E_1$ $X_2 = T+e+f_2+E_2$

Table 2.--Single adjective variance for the three formats under the appropriate assumptions of the four models.

MODELS	σ_X^2	Formats 1 Fixed adjective	2 Fixed concept multiple dimension	3 Fixed concept evaluative only	RANK* ORDER
	I	$\sigma_T^2 + \sigma_E^2$	$\sigma_T^2 + \sigma_E^2$	$\sigma_T^2 + \sigma_E^2$	1=2=3
	IIa Resampling	$\sigma_T^2 + \sigma_e^2 + \sigma_E^2$	$\sigma_T^2 + \sigma_e^2 + \sigma_E^2$	$\sigma_T^2 + \sigma_e^2 + \sigma_E^2$	1=2=3
	IIb	$\sigma_T^2 + \sigma_e^2 + \sigma_E^2$	$\sigma_T^2 + \sigma_e^2 + \sigma_E^2$	$\sigma_T^2 + \sigma_e^2 + \sigma_E^2$	1=2=3
	IIIa Resampling	$\sigma_T^2 + \sigma_e^2 + \sigma_E^2$	$\sigma_T^2 + \sigma_e^2 + \sigma_E^2$	$\sigma_T^2 + \sigma_e^2 + \sigma_\epsilon^2 + \sigma_E^2$	1=2<3
	IIIb	$\sigma_T^2 + \sigma_e^2 + \sigma_E^2$	$\sigma_T^2 + \sigma_e^2 + \sigma_E^2$	$\sigma_T^2 + \sigma_e^2 + \sigma_\epsilon^2 + \sigma_E^2$	1=2<3
	IVa Resampling	$\sigma_T^2 + \sigma_e^2 + \sigma_E^2$	$\sigma_T^2 + \sigma_e^2 + \sigma_E^2$	$\sigma_T^2 + \sigma_e^2 + \sigma_f^2 + \sigma_E^2$	1=2<3
	IVb	$\sigma_T^2 + \sigma_e^2 + \sigma_E^2$	$\sigma_T^2 + \sigma_e^2 + \sigma_f^2 + \sigma_E^2$	$\sigma_T^2 + \sigma_e^2 + \sigma_f^2 + \sigma_E^2$	1<2=3

*1 represents the Fixed adjective format, 2 represents the Fixed Concept, multiple dimension format and 3 represents the Fixed concept, evaluative only format.

Table 3.--Models which predict the rank orders of variances obtained.

Rank order of the variances	Models which predict this rank order
$1 = 2 = 3$	I, IIa, IIb
$1 = 2 < 3$	IIIa, IIIb, IVa
$1 < 2 = 3$	IVb

rank orders among the variances found in Table 2 are listed. Accompanying each relationship is the model or models which predict the rank order. From the relationships displayed in Table 3 only model IVb is clearly differentiated from the other models.

For each of the models, an expression for the correlation between two functionally synonymous adjectives was developed. These correlation coefficients for each of the three formats are displayed in Table 4. The rank order relationships among the correlations for the three formats predicted by each of the models are displayed in the last column of this table.

In this table, each of the models predicts a definite rank order among the correlations except model IIIb and model IVb. In model IIIb, the interadjective correlation for the fixed concept, multiple dimension format is predicted to be the largest, but the order of the remaining two formats is undetermined. Thus, there

Table 4.--Correlation coefficients for the three formats as developed in each of the models.

		Formats			
	$r_{X_1X_2}$	1 Fixed adjective	2 Fixed concept multiple dimension	3 Fixed concept evaluative only	Rank Order Relationship
MODELS	I	$\frac{\sigma_T^2}{\sigma_T^2 + \sigma_E^2}$	$\frac{\sigma_T^2}{\sigma_T^2 + \sigma_E^2}$	$\frac{\sigma_T^2}{\sigma_T^2 + \sigma_E^2}$	1 = 2 = 3
	IIa Resampling	$\frac{\sigma_T^2}{\sigma_T^2 + \sigma_e^2 + \sigma_E^2}$	$\frac{\sigma_T^2}{\sigma_T^2 + \sigma_e^2 + \sigma_E^2}$	$\frac{\sigma_T^2}{\sigma_T^2 + \sigma_e^2 + \sigma_E^2}$	1 = 2 = 3
	IIb	$\frac{\sigma_T^2}{\sigma_T^2 + \sigma_e^2 + \sigma_E^2}$	$\frac{\sigma_T^2 + \sigma_e^2}{\sigma_T^2 + \sigma_e^2 + \sigma_E^2}$	$\frac{\sigma_T^2 + \sigma_e^2}{\sigma_T^2 + \sigma_e^2 + \sigma_E^2}$	1 < 2 = 3
	IIIa Resampling	$\frac{\sigma_T^2}{\sigma_T^2 + \sigma_e^2 + \sigma_E^2}$	$\frac{\sigma_T^2}{\sigma_T^2 + \sigma_e^2 + \sigma_E^2}$	$\frac{\sigma_T^2}{\sigma_T^2 + \sigma_e^2 + \sigma_\epsilon^2 + \sigma_E^2}$	1 = 2 > 3

Table 4.--Continued.

MODELS	$r_{X_1X_2}$	Formats 1 Fixed adjective	2 Fixed concept multiple dimension	3 Fixed concept evaluative only	Rank Order Relationship
	IIIb	$\frac{\sigma_T^2}{\sigma_T^2 + \sigma_e^2 + \sigma_E^2}$	$\frac{\sigma_T^2 + \sigma_e^2}{\sigma_T^2 + \sigma_e^2 + \sigma_E^2}$	$\frac{\sigma_T^2 + \sigma_e^2}{\sigma_T^2 + \sigma_e^2 + \sigma_\epsilon^2 + \sigma_E^2}$	$1 < 3 < 2$ $1 = 3 \quad 2$ $3 < 1 < 2$
	IVa Resampling	$\frac{\sigma_T^2}{\sigma_T^2 + \sigma_e^2 + \sigma_E^2}$	$\frac{\sigma_T^2}{\sigma_T^2 + \sigma_e^2 + \sigma_E^2}$	$\frac{\sigma_T^2}{\sigma_T^2 + \sigma_e^2 + \sigma_f^2 + \sigma_E^2}$	$1 = 2 > 3$
	IVb	$\frac{\sigma_T^2}{\sigma_T^2 + \sigma_e^2 + \sigma_E^2}$	$\frac{\sigma_T^2 + \sigma_e^2}{\sigma_T^2 + \sigma_e^2 + \sigma_f^2 + \sigma_E^2}$	$\frac{\sigma_T^2 + \sigma_e^2}{\sigma_T^2 + \sigma_e^2 + \sigma_f^2 + \sigma_E^2}$	$1 < 2 = 3$ $1 = 2 = 3$ $1 = 3 < 1$

are three possibilities; $1 = 3$, $1 < 3$, $1 > 3$. For model IVb, the only relation which is determined is that the correlations for the two fixed concept formats should be the same. The value of $r_{X_1X_2}$ in the fixed adjective model might be less than, equal to, or greater than the value for the fixed concept formats.

In Table 5, the predicted rank orders of the $r_{X_1X_2}$'s found in Table 4 are listed with the models which predict each relationship. Models IIIb and IVb are differentiated from other models on the basis of $r_{X_1X_2}$.

Table 5.--Models which predict each of the obtained rank orders of the interadjective correlations.

Predicted rank order of the interadjective correlations	Models predicting the relationship
$1 = 2 = 3$	I, IIa, IVb
$1 = 2 > 3$	IIIa, IVa
$1 < 2 = 3$	IIb, IVb
$1 < 3 < 2$	IIIb
$1 = 3 < 2$	IIIb
$3 < 1 < 2$	IIIb
$1 > 3 = 2$	IVb

The statistics presented above both suffer a very serious defect: the value of σ_T^2 is present in each expression. The models say nothing about σ_T^2. Furthermore the value of σ_T^2 would vary from concept to concept. Thus it acts as a "nuisance" variable of considerable magnitude. For this reason, a statistic was sought that eliminates σ_T^2. The statistic found was the factor analytic "unique variance." That is, the covariance matrix of the adjectives has the form

$$\begin{matrix} \sigma_{XX} + \sigma_\mu^2 & \sigma_{XX} & \sigma_{XX} & \cdots \\ \sigma_{XX} & \sigma_{XX} + \sigma_\mu^2 & \sigma_{XX} & \cdots \\ \sigma_{XX} & \sigma_{XX} & \sigma_{XX} + \sigma_\mu^2 & \cdots \\ \vdots & \vdots & \vdots & \end{matrix}$$

where σ_μ^2 is the unique variance. In classical reliability, $\sigma_{X_1X_2} = \sigma_T^2$ and $\sigma_\mu^2 = \sigma_E^2$ and thus the uniqueness does not contain σ_T^2. In the models presented here, neither of these formulas hold in general. However σ_T^2 is always a component of $\sigma_{X_1X_2}$ and therefore never a component of σ_μ^2. The computational formula actually used corresponded to the identity

$$\sigma_\mu^2 = \sigma_X^2 - \sigma_{X_1X_2} = \sigma_X^2 (1-r_{XX})$$

The expressions for the unique variance for each format within each model are presented in Table 6.

Table 6.--Unique variances for the three formats corresponding to each model.

	$\sigma_X^2(1-r_{X_1X_2})$	Formats 1 Fixed adjective	2 Fixed concept multiple dimension	3 Fixed concept evaluative only	Rank Order Relationship
MODELS	I	σ_E^2	σ_E^2	σ_E^2	1 = 2 = 3
	IIa Resampling	$\sigma_e^2 + \sigma_E^2$	$\sigma_e^2 + \sigma_E^2$	$\sigma_e^2 + \sigma_E^2$	1 = 2 = 3
	IIb	$\sigma_e^2 + \sigma_E^2$	σ_E^2	σ_E^2	1 > 2 = 3
	IIIa Resampling	$\sigma_e^2 + \sigma_E^2$	$\sigma_e^2 + \sigma_E^2$	$\sigma_e^2 + \sigma_\epsilon^2 + \sigma_E^2$	1 = 2 < 3
	IIIb	$\sigma_e^2 + \sigma_E^2$	σ_E^2	$\sigma_\epsilon^2 + \sigma_E^2$	1 > 3 > 2 1 = 3 > 2 3 > 1 > 2
	IVa Resampling	$\sigma_e^2 + \sigma_E^2$	$\sigma_e^2 + \sigma_E^2$	$\sigma_e^2 + \sigma_f^2 + \sigma_E^2$	1 = 2 < 3
		$\sigma_e^2 + \sigma_E^2$	$\sigma_f^2 + \sigma_E^2$	$\sigma_f^2 + \sigma_E^2$	1 > 2 = 3 1 = 2 = 3 1 < 2 = 3

Again model IIIb and model IVb each produce three possible rank orders of the formats. In model IIIb, the only relation determined is that the unique variance for the fixed concept, multiple dimension format will be less than the unique variance of the other two formats. If $\sigma^2_\epsilon > \sigma^2_e$ then $3 > 1 > 2$, if $\sigma^2_\epsilon = \sigma^2_e$ then $1 = 3 > 2$ and if $\sigma^2_e > \sigma^2_\epsilon$ then $1 > 3 > 2$. For model IVb, the only relation determined is that the unique variance of the two fixed concept formats should be equal. If $\sigma^2_e = \sigma^2_f$ then $1 = 2 = 3$; if $\sigma^2_f > \sigma^2_e$ then $1 < 2 = 3$.

Table 7 presents the predicted rank orders of unique variances among the formats and the models which predict each relationship. The models which are singled out by the unique variance are models IIIb and IVb.

Differentiation of the models using all three statistics. If all three of the previous discussions are combined, the models are still not fully differentiated. Models I and IIa are equivalent for all three statistics. Models IIIa and IVa also have the same rank order on all three statistics. Model IIb is confused with only model IVb on the synonomous adjective correlation and unique variance, and is differentiated from IVb in its pattern of variances. Model IIIb has a unique pattern of synonomous adjective correlations and unique variances. Model IVb is identified by its pattern of variances. This is summarized in Table 8 which lists those models

Table 7.--Models which predict the possible rank orders of the unique variances.

Predicted rank order of the unique variances, $\sigma_X^2(1-r_{X_1X_2})$	Models predicting the relationship
1 = 2 = 3	I, IIa, IVa
1 > 2 = 3	IIb, IVb
1 = 2 < 3	IIIa, IVa
1 > 3 > 2	IIIb
1 = 3 > 2	IIIb
3 > 1 > 2	IIIb
1 < 2 = 3	IVb

Table 8.--Models which are singled out by some combination of variances, correlations and unique variances.

Models differentiated from all other	Combination of paramaters necessary for differentiation
IVb	σ_X^2
IIIb	$r_{X_1X_2}$ or $\sigma_X^2 (1-r_{X_1X_2})$
IIb	requires all three

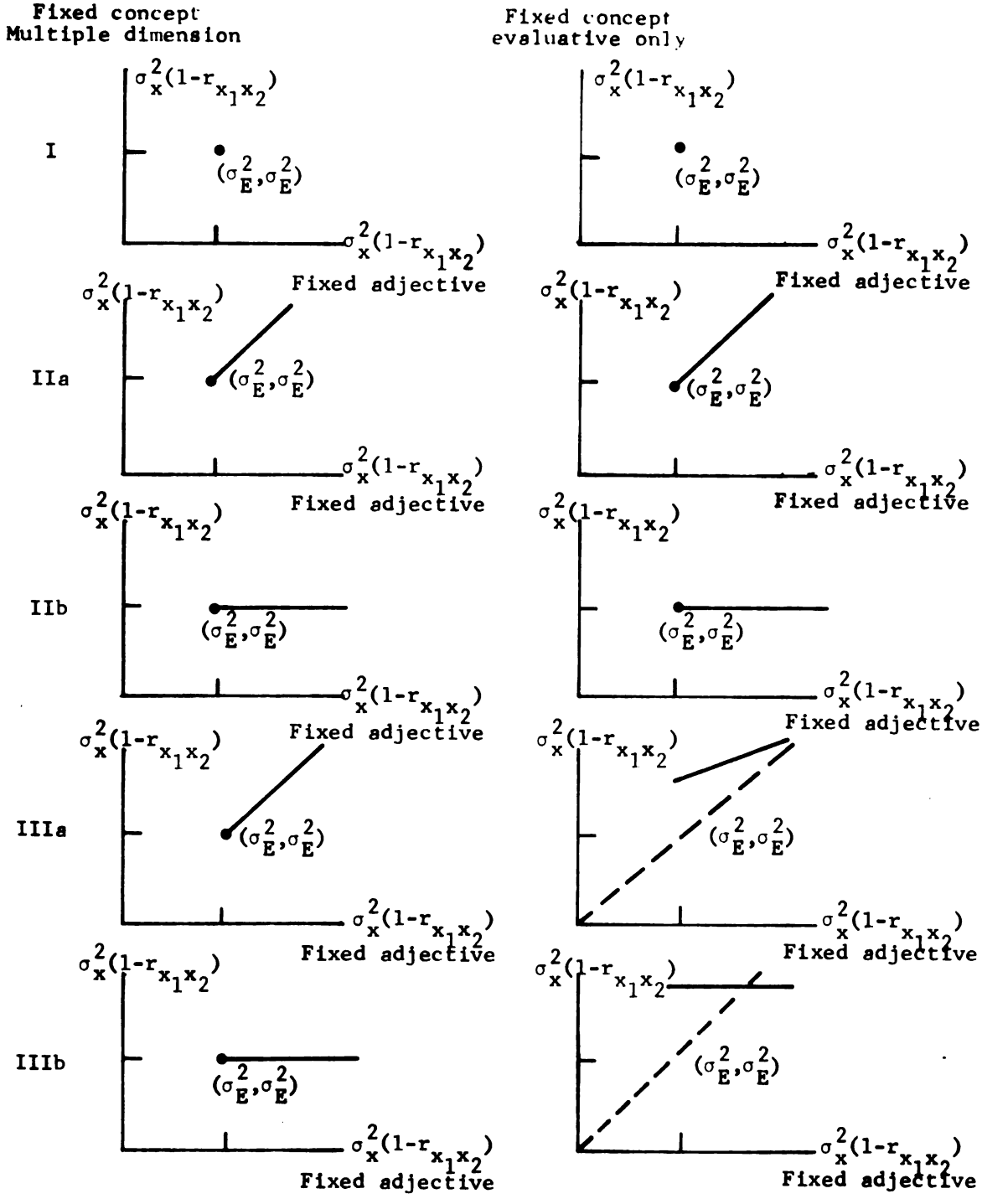
which are differentiated by some combinations of the three paramaters.

Differentiation of the models across concepts. If the models cannot be fully differentiated on a single concept, can they be differentiated if a set of concepts is considered? Here two ideas come to mind. First one might average each statistic across concepts. This has the advantage of eliminating (or greatly reducing) sampling error in the rank order comparisons discussed above for single concepts. And indeed the differentiation among models for these means is the same as the differentiation for single concepts.

A second method of using the data for a set of concepts that comes quickly to mind is to correlate values. Thus one might plot the fixed concept variances as a function of the fixed adjective variances. If you did, the disadvantage of the variances and synonomous adjective correlations would be immediately obvious. These statistics both contain the true score variance and the variance of true scores differs from concept to concept in ways and for reasons that are irrelevant to the models. On the other hand the unique variances have eliminated this component and yield much more meaningful comparisons.

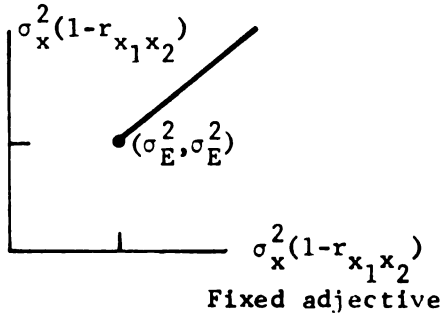
Figure 1 presents hypothetical graphs of the unique variances for each of the two fixed concept

Figure 1. Hypothetical graphs of the unique variances for a set of concepts from each of the two fixed concept conditions as a function of the unique variances of the set of concepts from the fixed adjective condition.


Fixed concept
Multiple dimension
Fixed concept
evaluative only
I
IIa
IIb
IIIa
IIIb
$\sigma_x^2(1-r_{x_1x_2})$
(σ_E^2,σ_E^2)
Fixed adjective

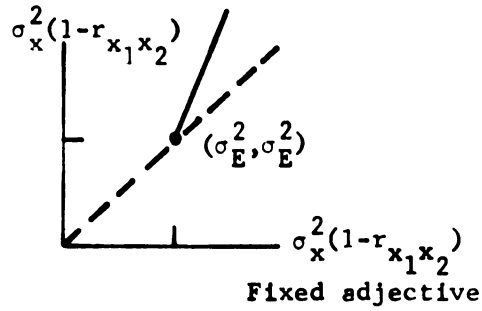
Fixed concept
Multiple dimension

IVa



Fixed adjective

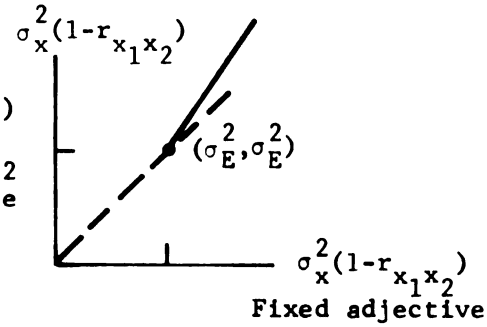
Fixed concept
evaluative only



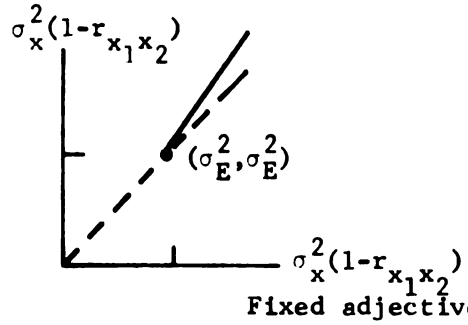
Fixed adjective

IVb (1)

$\sigma_f^2 > \sigma_e^2$



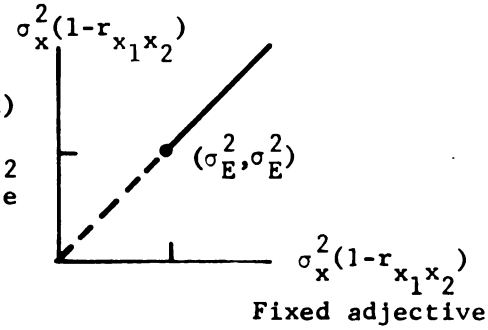
Fixed adjective



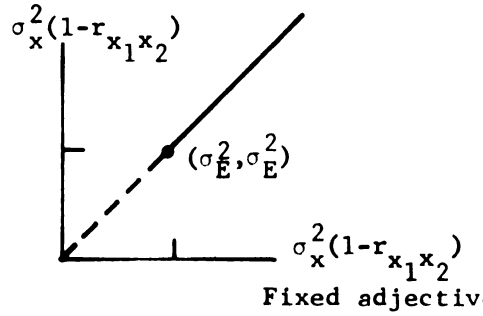
Fixed adjective

IVb (2)

$\sigma_f^2 = \sigma_e^2$



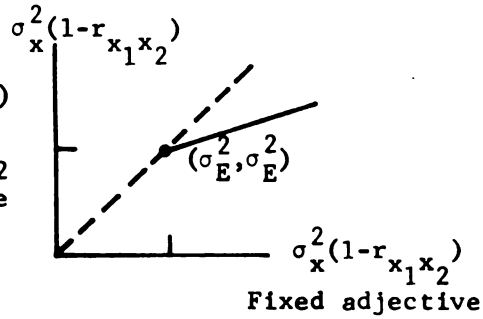
Fixed adjective



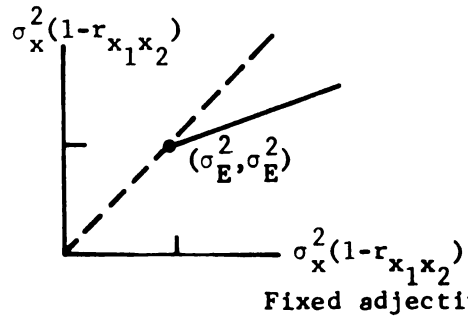
Fixed adjective

IVb (3)

$\sigma_f^2 < \sigma_e^2$



Fixed adjective



Fixed adjective

conditions as a function of the unique variance for the fixed adjective condition. The "purifying" assumptions required by these graphs will be given as the discussion progresses. First consider the graphs for model I. Since the process of converting a feeling to a pencil mark is the same for all concepts, the error introduced by this process, σ_E^2, is the same for all concepts and across conditions. Thus the entire graph consists of the single point (σ_E^2, σ_E^2). On the other hand, the error introduced by cognitive sampling varies with the cognitive complexity of the domain in question. Thus model IIa shows various points on the line y = X above the point (σ_E^2, σ_E^2). Model IIb shows various values greater than σ_E^2 on the X-axis while the y value is fixed at σ_E^2.

The second graph for model IIIa requires some explanation. The forced variance component, σ_ϵ^2, is never negative. Thus the graph is always above the line y = X. However the fact that the curve rejoins the line y = X reflects the assumption that σ_ϵ^2 and σ_e^2 are negatively correlated, i.e. the assumption that the greater the variance in the responses generated by heterogeneity in memory, the less the forced variance. The curve being a straight line reflects a completely gratuitous linearity assumption for this negative correlation. The graph for model IIIb dramatically shows the fact that in model IIIb the value of σ_E^2 is the same for all concepts.

The second graph for model IVa makes two "purifying"

assumptions. The fact that the distance from the line $y = X$ increases reflects the assumption σ_e^2 and σ_f^2 are both positively correlated with cognitive complexity and hence positively correlated with each other. The graph also assumes that both correlations are perfect and linear. Model IVb is complicated by the fact that σ_f^2 might be larger than, equal to, or less than σ_e^2. Hence three different graphs are shown.

The only two models which are not sharply differentiated by these graphs are models IIa and IVb. If $\sigma_f^2 = \sigma_e^2$, then model IVb yields exactly the same graph as model IIa.

THE EXPERIMENT

Instruments

This experiment was designed to evaluate the previously described models of semantic differential responses. Thus, three instruments were designed so that each matched one of the three formats discussed in the models. In order to make comparisons among the models, the same concepts and evaluative adjectives appeared in all three instruments. Each instrument consisted of 20 concepts rated on 5 evaluative scales. In addition, the fixed concept, multiple dimension and the fixed adjective instruments also contained five scales from other dimensions. The evaluative scales used in the instruments were; good-bad, productive-destructive, honest-dishonest, desirable-undesirable, valuable-worthless. The non-evaluative scale were; stodgy-inovative, declining-growing, exploitive-public spirited, inefficient-efficient, becoming more important-becoming less important. The concepts used in all of the instruments were; migrant workers (MW), draft (Dr), psychologists (Psy), open housing laws (Ohl), pollution (POLU), strikes (Str), computers (comp) law and order (L & O), interracial marriage (IM), disruptive protests

(DP), small businessmen (SBM), wire tapping (WT), labor unions (LU), large corporations (LC), integration by school bussing (Sch Bus), police (Poli), boycotts (Boy), civil rights movement (CRM), hippies (Hipp), use of sit-ins, lie-ins, etc. by civil right demonstrators (sit).

The fixed concept, multiple dimension format consisted of the twenty concepts, five evaluative scales plus the five scales from other dimensions. The order in which the scales were presented in this format alternated between evaluative and non-evaluative. Thus, all responses to an evaluative scale were separated by a response to non-evaluative scale.

The fixed concept, evaluative only format contained the same twenty concepts and five evaluative scales. The five non-evaluative scales were omitted from this format. In this format, evaluative responses to a given concept were given consecutively. The concepts were presented in the same order in both of the fixed concept formats.

The instrument which used the fixed-adjective format contained the same five evaluative and non-evaluative scales. The same twenty concepts were also used. In this format, a single bipolar scale was presented just preceding the list of twenty concepts and the concepts were always listed in the same order. For convenience in scoring, all three formats were printed

on IBM scoring forms. For the fixed concept, multiple dimension format, the first concept was presented with the ten adjective scales on which it was to be rated. Then the second concept was presented with the ten scales, etc. The adjectives in each scale were separated by seven response categories. The middle response category was distinguishable because it was printed in blue while the other categories were printed in red. The instrument which followed the fixed concept, evaluative only format had a similar spatial layout. The only difference was that five adjectives followed each concept.

The fixed adjective format was also printed on IBM forms. The current adjective scale was presented with seven response choices between each of the adjectives. The middle response category was again printed in blue while the others were in red. The twenty concepts were then listed below the current adjective scale.

To insure that the response to each concept was placed within the seven possible choices, the first letter of each of the bipolar adjectives was listed in line with each concept.

The instructions presented by Osgood (3) were modified to conform to both the IBM scoring procedure and the particular format being used. All three formats are presented in Appendix I.

Procedure and Subjects

The semantic differentials were administered to groups of about twenty subjects. All of the subjects within each group responded to the same format. The 119 subjects who responded to the fixed concept format multiple dimension format and the 78 subjects who responded to the fixed adjective format were students enrolled in an introductory psychology course at Michigan State University. Response to the fixed concept, evaluative only format were obtained from 91 subjects enrolled in a sophomore level psychology course. All subjects were volunteers who received credit which could be applied to their class grades. The data for the fixed concept, multiple dimension condition was generously provided by William J. Brown.

Results and Discussion

The four models previously described were concerned with the relationships among parallel adjective scales. So, the data analysis in this section will be restricted to adjective scales from a single semantic dimension, the evaluative dimension. The five evaluative scales and the twenty concepts combine to produce a total of 100 items from each of the three formats.

An examination of the 100 item correlation matrices for each of the three instruments indicated that all five evaluative scales acted as parallel items. That is,

within each concept, the adjective means and variances were about equal and the ten inter evaluative scales correlations for a given concept were all about the same size. Furthermore, all five evaluative adjectives on a given concept tended to display the same pattern of correlations with adjectives on the other concepts. The means and standard deviations of the adjective variances, means and inter correlations for each concept are presented in Appendix II.

Since the items are all assumed to have the same variance, that common variance was estimated by averaging the five obtained evaluative scale variances for each concept. Thus, for each format, twenty parallel item variances were produced. In the population the interadjective correlations for a given concept should also all be equal. For each concept, the interadjective correlation was estimated by averaging the ten interscale correlations. Thus, for each of the three instruments, estimates of the twenty parallel item correlations were produced. The third statistic to be estimated is the unique variance, $\sigma_X^2 (1-r_{X_1X_2})$. The formula used was $\overline{\sigma_X^2} (1-\overline{r}_{XX})$ where $\overline{\sigma_X^2}$ and $\overline{r}_{X_1X_2}$ are the averages just described. Thus within each format, the twenty concepts are characterized by three numbers, an estimated variance, an estimated interadjective correlation, and an

estimated unique variance. These three tables are presented in Appendix II.

The twenty values of each statistic were averaged across concepts to produce the means displayed in Table 9. The first three columns of the table are the mean values for the three formats. The fourth column gives the rank order relationships among the formats based on each of the three statistics. The final column of this table indicates which of the models would have predicted the obtained rank ordering.

Only one model was consistent with the rank order of the variances: model IVb. Two models predicted the rank order of the interadjective correlations: models IIIa and IVa.

None of the models predicted the relationship found for the unique variances. However, model IIIa and model IVa predict that the unique variances will order the formats as $1 = 2 < 3$. The unique variance values for the fixed adjective format, 1, is .72 and the fixed concept multiple dimension format, 2, is .79 which are fairly close to each other and "far" from the .98 of the third format. Thus, models IIIa and IVa come close to fitting the unique variance portion of the data.

Strictly speaking there was no model that fit the data for even two of the three statistics. However to the extent that models IIIa and IVa "come close" for

Table 9.--Mean values of the variance, interadjective correlation and unique variances.

	Instruments				
	1 Fixed adjective	2 Fixed concept multiple dimension	3 Fixed concept evaluative only	Rank Order	Models that fit
$\overline{\sigma_X^2}$	1.776	2.006	2.019	1<2=3	IVb
$\overline{r_{X_1X_2}}$	.566	.578	.490	1=2>3	IIIa IVa
$\overline{\sigma_X^2}\ (1-\overline{r}_{X_1X_2})$	.719	.792	.977	1<2<3	none

the unique variances, they would fit two statistics since they fit the interadjective correlation rank order. For variances these models predict 1 = 2 < 3 where the obtained was 1 < 2 = 3 which is at least in the right ball park.

To differentiate the models across concepts, the fixed concept unique variances are plotted as a function of the unique variance for the fixed adjective condition. Figure 2 presents the unique variances obtained in the fixed concept multiple dimension condition as a function of the unique variances obtained from the fixed adjective condition. The most glaring feature of this scatterplot is that "pollution" acts as an outlier. Therefore the two regression lines shown were calculated with the data for pollution deleted. The difference between the lines reflects the fact that even with sample sizes near 100, the sampling error in these variance estimates is large enough to greatly reduce the correlation. Furthermore the ambiguity introduced is serious. The upper regression line is consistent with model IVb (1), semantic over-differentiation with $\sigma_f^2 > \sigma_e^2$. The lower regression line fits none of the models. It is worth noting that the small slope in the lower regression line is largely the result of the single data point for "migrant workers."

Figure 3 presents the unique variances obtained from the fixed concept, evaluative only condition plotted as a function of the unique variances obtained under the

Figure 2. The unique variances from the fixed concept, multiple dimension condition plotted as a function of the unique variances from the fixed adjective condition along with both regression lines.

*Calculated without the data for pollution.

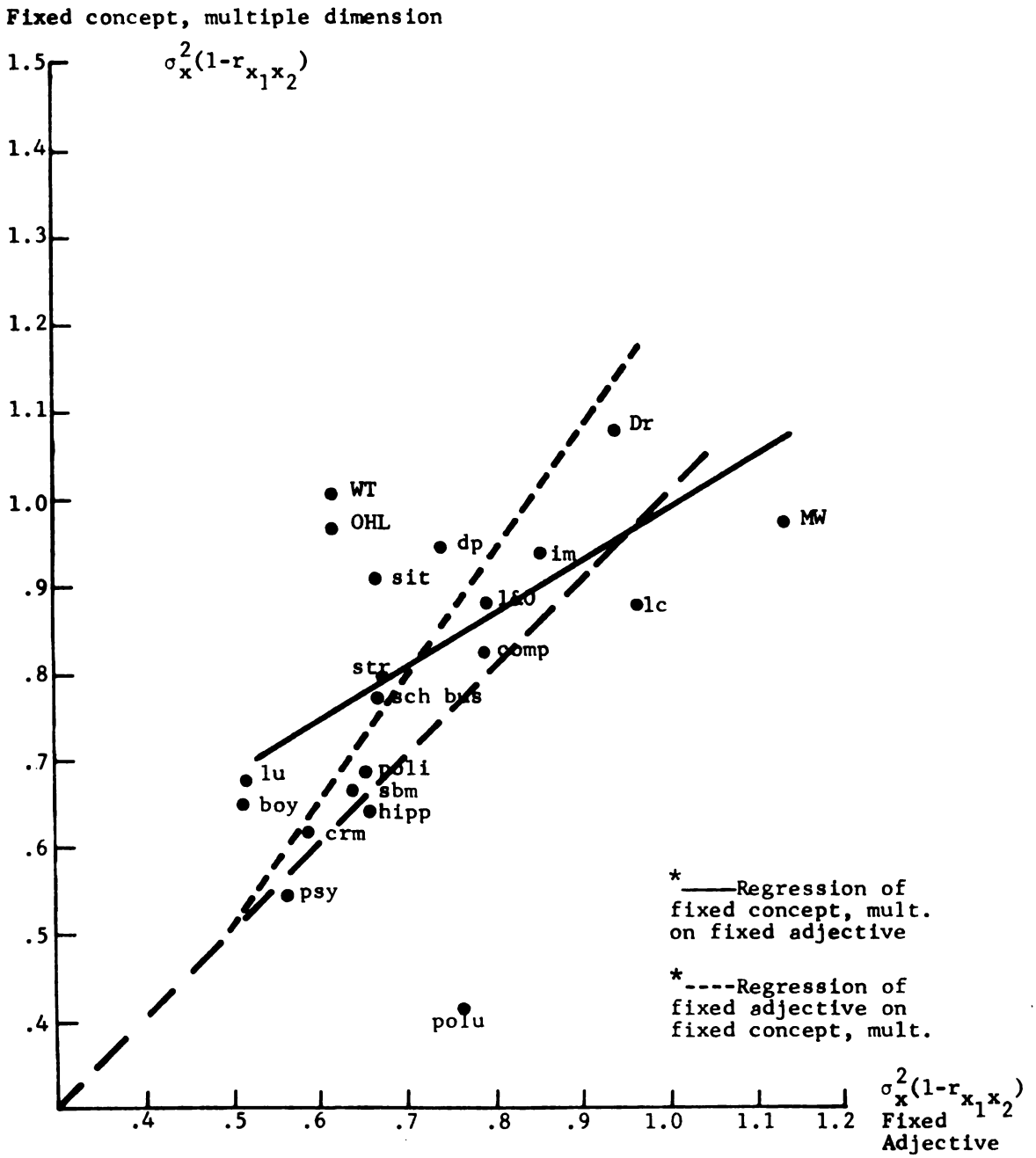
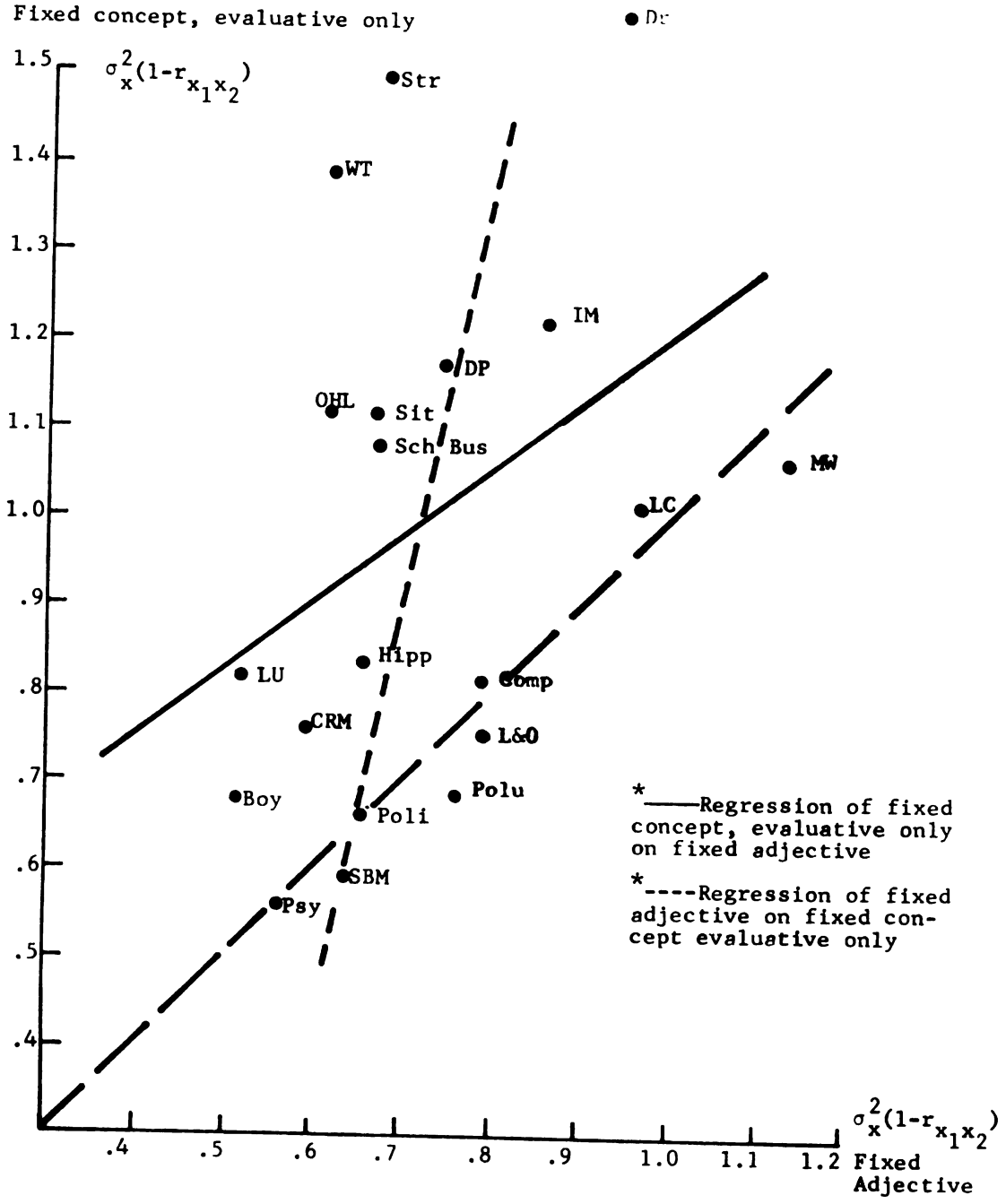

Fixed concept, multiple dimension
$\sigma_x^2(1-r_{x_1x_2})$
1.5
1.4
1.3
1.2
1.1
1.0
.9
.8
.7
.6
.5
.4
WT
OHL
Dr
MW
dp
im
sit
lc
comp
str
sch bks
lu
poli
sbm
boy
hipp
crm
psy
polu
*——Regression of fixed concept, mult. on fixed adjective
*----Regression of fixed adjective on fixed concept, mult.
.4
.5
.6
.7
.8
.9
1.0
1.1
1.2
$\sigma_x^2(1-r_{x_1x_2})$ Fixed Adjective

Figure 3. The unique variances from the fixed concept, evaluative only condition plotted as a function of the unique variances from the fixed adjective condition and both regression lines.

*Calculated without the data for pollution.


Fixed concept, evaluative only
$\sigma_x^2(1-r_{x_1x_2})$
1.5
1.4
1.3
1.2
1.1
1.0
.9
.8
.7
.6
.5
.4
Str
WT
IM
DP
OHL
Sit
Sch
Bus
MW
LC
Hipp
LU
Comp
CRM
L&O
Boy
Polu
Poli
SBM
Psy
*——Regression of fixed concept, evaluative only on fixed adjective
*----Regression of fixed adjective on fixed concept evaluative only
.4
.5
.6
.7
.8
.9
1.0
1.1
1.2
$\sigma_x^2(1-r_{x_1x_2})$ Fixed Adjective

fixed adjective condition. As in the previous scatterplot, the models do not dictate the order in which the variables should be plotted. Thus either regression line can be considered. Again sampling error is sufficient to produce ambiguity in the results. The upper regression line is consistent with models IVa and IVb (1), the semantic overdifferentiation models. The lower regression line would technically be consistent with model IIIa, the forced variance model. However, the slope is much too high for the substantive assumption. This becomes crystal clear when stated in terms of the point where the regression line crosses the line y - x (a point well beyond the limits of the graph). According to the model, it is not until this point that the "natural" variance in the subject's responses terminates the forced variance process. Which of the two regression lines is nearer the truth? There are two points to be considered. First, sampling error, like unreliability, produces the greatest deviation in the regression line if the variable with the lower "reliability" is plotted on the X-axis. Since sampling error is somewhat greater for the fixed adjective condition (N = 78 vs. N = 91), while the spread is much greater for the variances in the evaluative only condition, the fallibility of the fixed adjective variances is much larger. Thus it is the upper regression line which is closer to the line without sampling error. Second, it is worth noting that the principal reason for

the deviation of the two lines is the single data point for migrant workers. Furthermore it will be recalled that it was the migrant workers point which was off in the previous graph. In fact both points are off in exactly the same way: either too "low" or too far to the right. Since the X-coordinate is the same number in both graphs, it is tempting to assume that this single number is about 20 percent too large by sampling error. This would largely eliminate the problems in interpreting both graphs.

In principle, it is possible to correct the regression lines in Figures 2 and 3 to remove the effect of the sampling error in the uniquenesses. If the unique variance were 3.0 and it was based on a sample of 101 subjects, then its standard error would be

$$\frac{2.3^2}{101-1} = \frac{3}{10}\,2 = .423$$

if the distribution of responses was approximately normal. The details of this procedure are spelled out in Appendix 3. To the extent that this procedure is valid, the scatter plots in Figures 4 and 5 are drawn with the regression lines that would have been found had there been no sampling error in the uniquenesses.

If the regression lines displayed in both Figure 4 and Figure 5 are compared to each other and to the hypothetical graphs in Figure 1, then none of the models matches the data. The hypothetical graphs for model

Figure 4. The unique variances from the fixed concept multiple dimension condition plotted as a function of the unique variances from the fixed adjective condition and the regression of the fixed concept, multiple dimension condition on the fixed adjective condition corrected for attenuation.

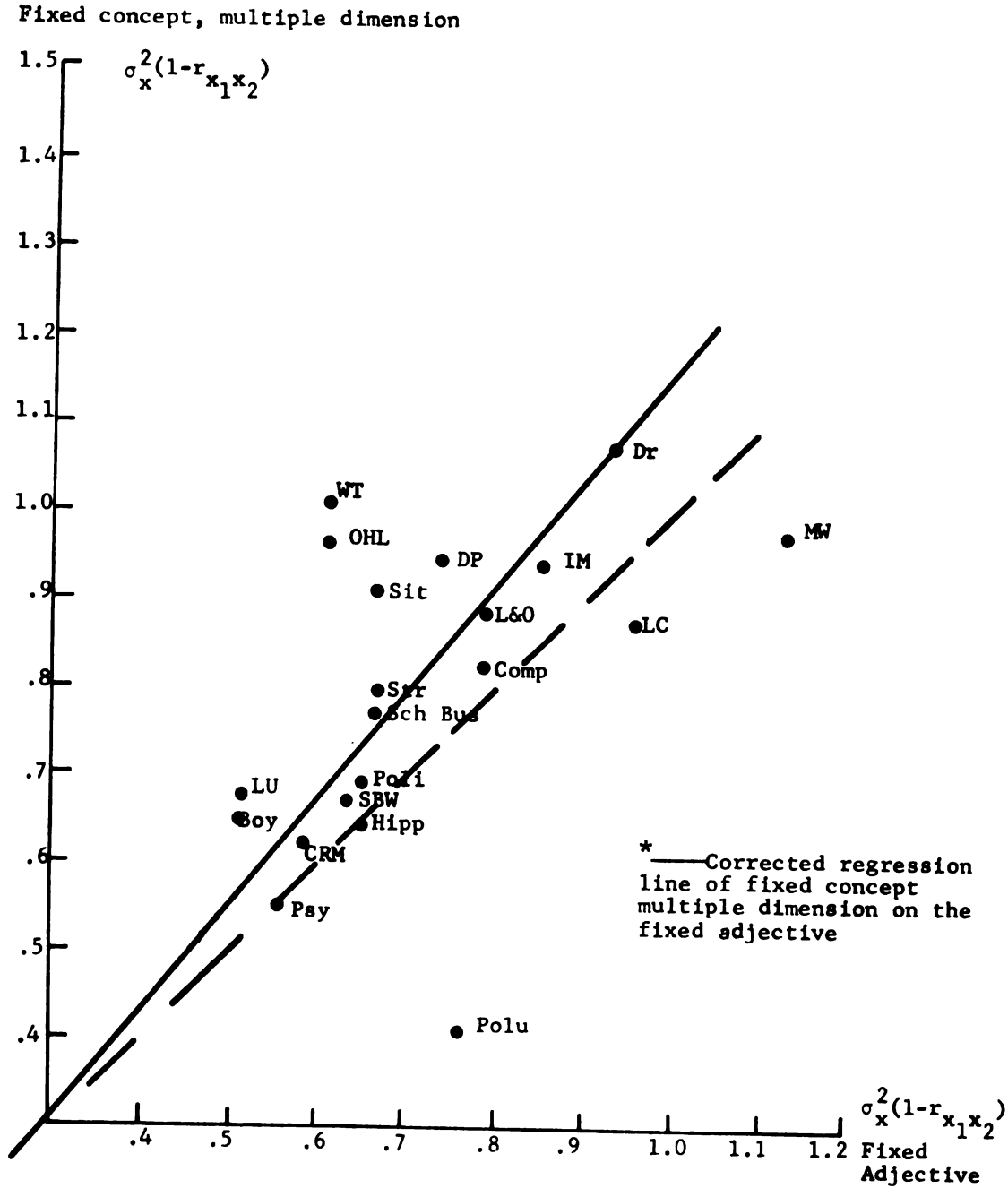
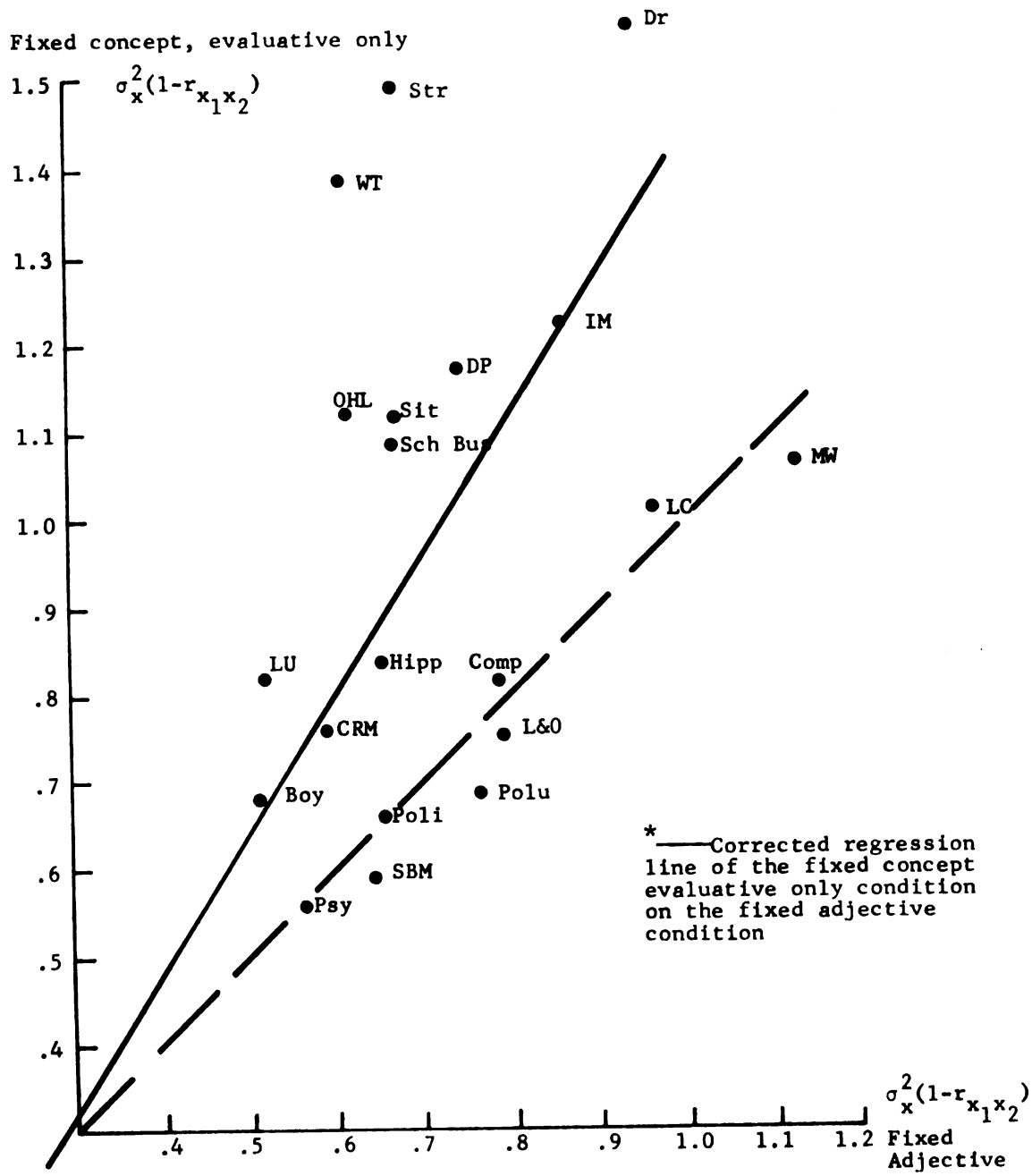

Fixed concept, multiple dimension
$\sigma_x^2(1-r_{x_1x_2})$
1.5
1.4
1.3
1.2
1.1
1.0
.9
.8
.7
.6
.5
.4
Dr
WT
OHL
MW
DP
IM
Sit
L&O
LC
Comp
Str
Sch Bus
LU
Poli
SBW
Boy
Hipp
CRM
Psy
Polu
*——Corrected regression line of fixed concept multiple dimension on the fixed adjective
.4
.5
.6
.7
.8
.9
1.0
1.1
1.2
$\sigma_x^2(1-r_{x_1x_2})$ Fixed Adjective

Figure 5. The unique variances from the fixed concept, evaluative only condition plotted as a function of the unique variances from the fixed adjective condition and the regression of the fixed concept evaluative only on the fixed adjective condition corrected for attenuation.


Fixed concept, evaluative only
$\sigma_x^2(1-r_{x_1x_2})$
1.5
1.4
1.3
1.2
1.1
1.0
.9
.8
.7
.6
.5
.4
Dr
Str
WT
IM
DP
OHL
Sit
Sch Bu
MW
LC
LU
Hipp
Comp
CRM
L&O
Boy
Polu
Poli
SBM
Psy
*——Corrected regression line of the fixed concept evaluative only condition on the fixed adjective condition
.4
.5
.6
.7
.8
.9
1.0
1.1
1.2
$\sigma_x^2(1-r_{x_1x_2})$ Fixed Adjective

IVb(1) are similar to those obtained. However, the obtained regression lines should be the same for model IVb(1) and they are not. On the other hand, model IVa predicts a difference between the two which is in the right direction (graph 2 > graph 1), but it gives poor absolute fit to the first graph. This suggests that a hypothetical graph that gives good fit could be obtained by geometrically averaging the predicted graphs for models IVa and IVb(1). Figure 6 displays hypothetical graphs of the unique variances for each of the two fixed concept conditions as a function of the unique variance for the fixed adjective condition under this averaging model.

Conglomerate Model

One possible justification for this procedure is to suppose that half the subjects resample their cognitive domains prior to each response in all conditions while the other half of the subjects do not resample. If half the subjects resample and half don't, then the resulting effect on the unique variances would be to average the two expressions obtained for models IVa and IVb(1). Since the fixed adjective equations are the same for both models, this means that it is the fixed concept expressions which change, i.e. the y-values of the graphs are averaged as claimed.

This conglomerate model will also make predictions

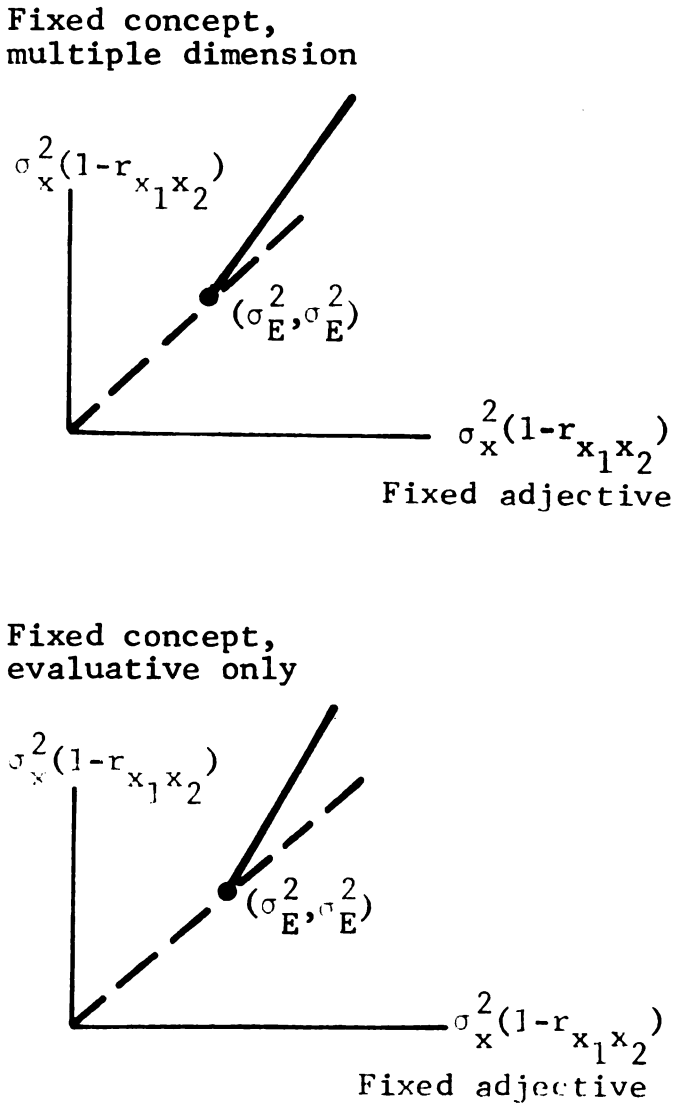


Figure 6. Hypothetical graphs of the unique variances of the two fixed concept conditions as a function of the unique variances for the fixed adjective condition for the conglomerate model.

for variances and interiten correlations. Table 10 gives the expression for σ_X^2, $r_{X_1X_2}$ and $\sigma_X^2 (1-r_{X_1X_2})$ for this model. As might be expected, it correctly predicts the rank order of the unique variances, $1 < 2 < 3$. However, it does not correctly predict the rank order of the variances, $1 > 2 = 3$. The unique variance prediction that $\sigma_3^2-\sigma_2^2 = \sigma_2^2-\sigma_1^2 = \frac{\sigma_f^2}{2}$ is parameter free. The rank order of the interadjective correlations is indeterminate. However the observed rank order, $1 = 2 > 3$, is obtained if

$$\frac{\sigma_f^2}{\sigma_e^2} = \frac{\sigma_T^2 + \sigma_e^2 + \sigma_E^2}{\sigma_T^2}$$

The principal part of this assumption is that $\sigma_f^2 > \sigma_e^2$, and this has already been assumed.

Thus the difficulty in accepting the conglomerate model boils down to the fact that the two fixed concept variances are equal. The three variances are 1.78 for the fixed adjective format, 2.01 for the fixed concept, multiple dimension format and 2.02 for the fixed concept, evaluative only format. If the variance for the fixed concept, multiple dimension format were 1.90 instead of 2.01 (a difference of 5 percent), the model would fit perfectly. Since these numbers are averages of twenty variances based on about 100 subjects each, sampling error can be eliminated as an explanation. The next most likely candidate for a "small" nuisance variable is

Table 10.--Expression for σ_X^2, $r_{X_1X_2}$ and $\sigma_X^2(1-r_{X_1X_2})$ for the conglomerate model.

	Formats			
	1 Fixed adjective	2 Fixed concept multiple dimension	3 Fixed concept evaluative only	Rank Order
σ_X^2	$\sigma_T^2 + \sigma_e^2 + \sigma_E^2$	$\sigma_T^2 + \sigma_e^2 + \sigma_f^2 + \sigma_E^2$	$\sigma_T^2 + \sigma_e^2 + \sigma_f^2 + \sigma_E^2$	$1 < 2 < 3$
$r_{X_1X_2}$	$\frac{\sigma_T^2}{\sigma_T^2 + \sigma_e^2 + \sigma_E^2}$	$\frac{\sigma_T^2 + \sigma_e^2/2}{\sigma_T^2 + \sigma_e^2 + \sigma_f^2 + \sigma_E^2}$	$\frac{\sigma_T^2 + \sigma_e^2/2}{\sigma_T^2 + \sigma_e^2 + \sigma_f^2 + \sigma_E^2}$	indeter- minate
$\sigma_X^2(1-r_{X_1X_2})$	$\sigma_e^2 + \sigma_E^2$	$\frac{\sigma_e^2}{2} + \frac{\sigma_f^2}{2} + \sigma_E^2$	$\frac{\sigma_e^2}{2} + \sigma_f^2 + \sigma_E^2$	$1 < 2 < 3$

population inhomogeneity. And this could be a factor in the present case. The multidimensional format was given to a freshman introductory psychology course while the evaluative only format was given to a sophomore introductory experimental psychology course. If the average standard deviation was 2½ percent larger for the freshman, the results would be explained. It should be noted that this explanation puts the inhomogeneity in σ_T^2. Thus the unique variances (means or plots) would be uneffected. The interadjective correlations would be affected by this assumption, but not greatly.

Implications of the model. If the conglomerate model holds, what implications can be drawn for the reliability of the semantic differential? The key fact about the conglomerate model is that it is a "correlated errors" model. For the fixed adjective condition, the errors are independent and $r_{XX} = r_{XT}^2$. Thus in the fixed adjective condition it is proper to use coefficient alpha to correct for attenuation, test for no change, etc. But in the fixed concept conditions, this is not true. In the fixed concept conditions, the "errors" are correlated i.e.

$$r_{X_1X_2} = \frac{\sigma_T^2 + \sigma_e^2/2}{\sigma_X^2}$$

while

$$r_{XT}^2 = \frac{\sigma_T^2}{\sigma_X^2}$$

Now r_{XT}^2 is the proper coefficient for correction for attenuation, but it is r_{XX} that is estimated by the usual reliability formulas (whether as explicitly as computing $\overline{r_{X_1X_2}}$, or implicitly in coefficient alpha). In this sense r_{XX} can be regarded as "spuriously" high for this model.

How big is the error? That is, by how much is r_{XX} larger than r_{XT}^2 ? A quick answer is given by reversing the equation for r_{XX}

$$r_{XX} = r_{XT}^2 + \frac{\sigma_e^2/2}{\sigma_X^2}$$

to yield

$$r_{XT}^2 = r_{XX} - \frac{\sigma_e^2/2}{\sigma_X^2}$$

However an estimate of σ_e^2 is required to use this formula. A reasonable place to look for such an estimate is in the unqiue variances. After fiddling with Table 6 for a while, some version of the following identity emerges. Let u_1, u_2, and u_3 be the unique variances for the three formats. Then

$$u_3 - 2u_2 + u_1 = \frac{\sigma_e^2}{2}$$

Thus for a typical concept in the present data, the estimate is

$$\frac{\sigma_e^2}{2} = .977 - 2(.792) + .719 = .112$$

The "proper" reliability for the fixed concept, multiple

dimension format is then given by

$$r^2_{XT} = r_{XX} - \frac{\sigma^2_e/2}{\sigma^2_X} = .578 - \frac{.112}{2.006} = .522$$

Thus for this format r_{XX} is "off" by .056 or a little better than 10 percent too high. If these values are extended by the Spearman-Brown formula to yield the reliability for a 5-adjective evaluative score,

$$\text{alpha} = r_{\overline{XX}} = \frac{5(.578)}{4(.578)+1} = .87$$

while the "proper" reliability is

$$r^2_{XT} = \frac{5(.522)}{4(.522)+1} = .84$$

Thus coefficient alpha is .03 too high or off by about 3½ percent.

The preceeding points are spelled out by a larger number of examples in Table 11. The first three columns of the table give assorted statistics for each of the three formats for a typical concept in the present data. The last column indicates what would happen if either lower reliability adjectives (such as clean-dirty, sick-well, etc.) were used or the concepts were choosen so that the feelings of the subject population were more in concordance (e.g. pollution in the present data). In the first two rows of the table, the "spuriousness" of the reliability of a single adjective varies from 0 for the fixed adjective format to 13 percent for the evaluative only, fixed concept format to 19 percent for the hypothetical low reliability concept. The rapid drop

Table 11.--Reliability data for a typical concept from each of the three instruments and for a concept rated on hypothetical low reliability adjectives.

	Fixed adjective	Fixed concept multiple dimension	Fixed concept evaluative only	Hypothetical low reliability adjectives
$r_{X_1X_2}$	.566	.578	.490	.350
r^2_{XT}	.566	.522	.434	.294
$r_{\overline{X_1X_2}}$	.87	.87	.83	.73
$r^2_{\overline{X}T}$	.87	.84	.79	.68
$\sqrt{r_{\overline{X}_1\overline{X}_2}}$	.93	.93	.91	.85
$r^2_{\overline{X}T}$	.93	.92	.89	.82

off becomes more rapid as the reliability goes down. If a five adjective evaluative scale is used, the Spearman-Brown formula produces the entries in the second pair of rows in Table 11. Here the spuriousness ranges from 0 for the fixed adjective format to five percent for the evaluative only to seven percent for the hypothetical low reliability concept. These last reliabilities are suitable for consideration in a context where two scales are to be corrected for attenuation (or change over time). If only one variable is to be corrected, for attenuation the "one-sided" correction formula will apply, and the appropriate coefficients are the square roots of the corresponding "two-sided" coefficients. These are given in the third pair of rows in Table 11. Here the error varies from 0 to 2 percent to 4 percent for the low reliability concept.

Thus a standard fixed concept, multidimensional five adjective instrument had a conventionally estimated reliability that was spuriously high by only 5 percent. This is not intolerable in most contexts, but if the population homogeneity goes up by just a little bit, the error goes to 7 percent or more. At this point "correction for attenuation" introduces a serious degree of error.

SUMMARY

This paper has been concerned with the internal consistency of the semantic differential. The initial hypothesis was that the usual method of presenting all of the adjectives for a single concept at the same time might lead to reliability data with "correlated errors." In grappling with the data a series of four formal mathematical models were created and tested against data. The model which fit best assumed (1) that subjects draw only a sample of their belief system in responding to the adjectives, (2) that the usual semantic differential format causes subjects to create idiosyncratic definitions which "overdifferintiate" the evaluative adjectives, and (3) that some of the subjects draw only one cognitive sample to make all the responses while others draw a new sample to make each response. The data collected were interpreted as showing a definite correlated errors component in the usual semantic differential format. The discussion noted that for high population homogeneity on a concept, this could result in a very substantial "spuriousness" in reliability estimates using any measure of internal consistency.

LIST OF REFERENCES

LIST OF REFERENCES

1. Stagmer, R. and C. E. Osgood. Impact of War on a Nationalistic Form of Reference: Changes in General Approval and Qualitative Patternings of Certain Stereotypes. J. of Soc. Psych., 1946, 24, 187-215.

2. Osgood, C. E. The Nature and Measurement of Meaning. Psychological Bulletin, 1952, 49(3), 197-237.

3. Osgood, C. E., G. J. Suci, and P. H. Tannebaum. The Measurement of Meaning. Urbana, Ill.: The University of Illinois Press, 1957.

4. Wells, W. D. and G. Smith. Four Semantic Rating Scales Compared. J. of Applied Psych., 1960, 44(6), 393-397.

5. Jenkins, J. J., W. A. Russel, and G. J. Suci. An Atlas of Semantic Profiles for 360 Words. Amer. J. Psychol., 1958, 71, 688-699.

6. Miran, M. S. The Influence of Instruction Modification Upon Test-Retest Reliabilities of the Semantic Differential. Ed. and Psych. Measurement, 1961, 21(4), 883-893.

7. DiVesta, F. J. and W. Dick. The Test-Retest Reliability of Children's Ratings on the Semantic Differential. Ed. and Psych. Measurement,

8. Block, J. An Unprofitable Application of the Semantic Differential. J. of Consulting Psych., 1958, 22(3), 235-236.

9. Spearman, C. The Proof and Measurement of Association Between Two Things. Amer. J. Psych., 1904, 15, 72-101.

10. Cronbach, L. J. Coefficient Alpha and the Internal Structure of Tests. Psychometrika, 1951, 16, 297-334.

APPENDICES

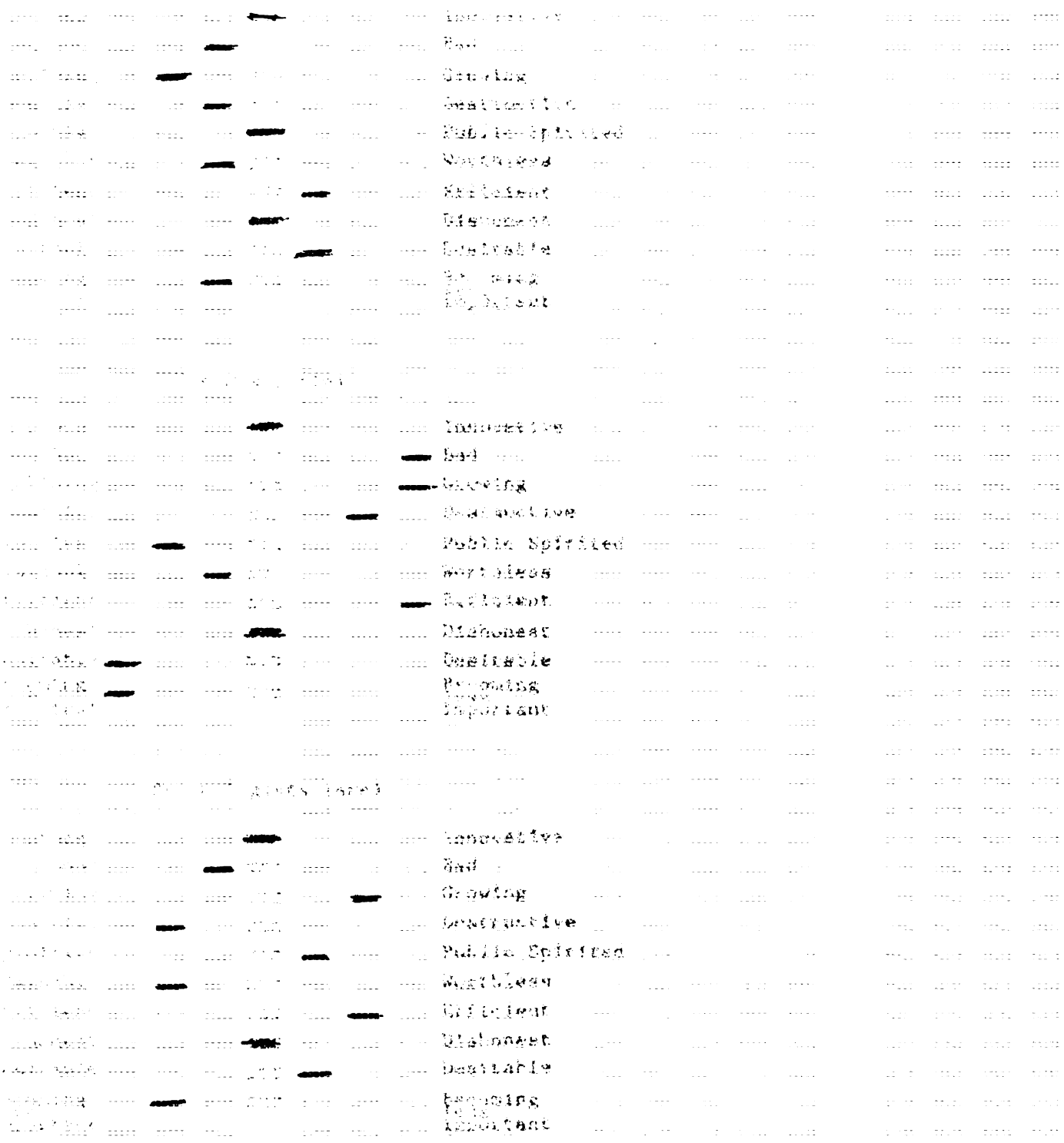
APPENDIX I

APPENDIX I

This appendix contains one sample page from each of the three semantic differentials used in the experiment. The first page is an example of fixed adjective format, the second an example of the fixed concept, multiple dimension format and the third is an example of the fixed concept, evaluative only format.

Symbolizes		Words Spelled
Migrant Workers(are)	R	I
Draft(is)	S	J
Psychologists(are)	S	P
Openhousing Laws(are)	S	P
Pollution(is)	S	P
Strikes(are)	S	P
Computers(are)	S	P
Law and Order(is)	S	P
Interracial Marriage(is)	S	P
Disruptive Protests(are)	S	P
Small Business Man(are)	S	P
Wire tapping(is)	S	P
Labor Unions(are)	S	P
Large Corporations(are)	S	P
Segregation by School Busing	S	P
Police(are)	S	P
Boycotts(are)	S	P
Civil Rights Movement(is)	S	P
Hippies(are)	S	P
Use of sit-in, lie-in etc. by Civil Rights Demonstrators(is)	S	P

Valuable		Worthless
Migrant Workers(are)	V	W
Draft(is)	V	W
Psychologists(are)	V	W
Openhousing Laws(are)	V	W
Pollution(is)	V	W
Strikes(are)	V	W
Computers(are)	V	W
Law and Order(is)	V	W
Interracial Marriage(is)	V	W
Disruptive Protests(are)	V	W
Small Business Man(are)	V	W
Wire tapping(is)	V	W
Labor Unions(are)	V	W
Large Corporations(are)	V	W
Segregation by School Busing	V	W
Police(are)	V	W
Boycotts(are)	V	W
Civil Rights Movement(is)	V	W
Hippies(are)	V	W
Use of sit-in, lie-in etc. by Civil Rights Demonstrators(is)	V	W



APPENDIX II

APPENDIX II

The following three tables contain the values of $\bar{\sigma}_X^2$, $\bar{r}_{XX}$, and $\bar{\sigma}_X^2(1-\bar{r}_{XX})$ for each of the twenty concepts in the three semantic differential formats. The values in Table 12 for $\overline{\sigma_X^2}$ were obtained by averaging the σ_X^2's for the five evaluative scales for each concept. The values for $\bar{r}_{XX}$ in Table 13 were computed by averaging the ten into evaluative scale correlations for each concept. Finally, the values of $\overline{\sigma_X^2}(1-\bar{r}_{XX})$ in Table 14 were obtained by using the averages presented in Tables 12 and 13.

Table 12.--Means and standard deviations of $\bar{\bar{X}}$ for each concept across the five evaluative adjectives.

	Condition					
	Fixed adjective		Fixed concept multiple dimension		Fixed concept evaluative only	
	$\bar{\bar{X}}$	$\sigma_{\bar{X}}$	$\bar{\bar{X}}$	$\sigma_{\bar{X}}$	$\bar{\bar{X}}$	$\sigma_{\bar{X}}$
Migrant Workers	3.19	.358	3.33	.375	3.28	.541
Draft	5.54	.315	5.36	.450	5.32	.432
Psychologists	2.41	.348	2.20	.273	2.19	.404
Open Housing Laws	2.56	.411	2.60	.383	2.76	.400
Pollution	6.49	.605	6.52	.726	6.53	.636
Strikes	3.70	.340	3.59	.303	3.79	.448
Computers	2.49	.304	2.04	.330	1.88	.411
Law and Order	2.82	.380	2.77	.596	2.21	.714
Interracial Marriage	3.26	.202	3.49	.248	3.63	.353
Disruptive Protests	4.53	.254	4.82	.344	5.13	.457
Small Businessmen	3.06	.246	2.64	.395	2.50	.453
Wire Tapping	5.27	.327	5.22	.743	5.14	.898
Labor Unions	3.43	.318	3.08	.470	2.94	.658
Large Corporations	3.95	.440	3.28	.560	3.02	.923

Table 12.--Continued.

	Condition					
	Fixed adjective		Fixed concept multiple dimension		Fixed concept evaluative only	
	$\bar{\bar{X}}$	$\sigma_{\bar{X}}$	$\bar{\bar{X}}$	$\sigma_{\bar{X}}$	$\bar{\bar{X}}$	$\sigma_{\bar{X}}$
Integration by school bussing	4.37	.153	4.40	.286	4.69	.376
Police	3.07	.338	2.83	.378	2.57	.466
Boycotts	3.21	.212	3.23	.154	3.30	.300
Civil Rights Movement	2.43	.231	2.26	.336	2.34	.472
Hippies	3.69	.224	3.43	.325	3.56	.263
Use of sit-ins, lie-ins, etc. by Civil Rights Demonstrators	3.23	.086	3.41	.232	3.49	.264

Table 13.--Means and standard deviations of σ_X^2 within each concept, across the five evaluative scales.

	Condition					
	Fixed adjective		Fixed concept multiple dimension		Fixed concept evaluative only	
	$\overline{\sigma_X^2}$	$\sigma_{\sigma_X^2}$	$\overline{\sigma_X^2}$	$\sigma_{\sigma_X^2}$	$\overline{\sigma_X^2}$	$\sigma_{\sigma_X^2}$
Migrant Workers	1.858	.362	1.700	.200	1.471	.419
Draft	2.282	.366	2.449	.322	2.815	.473
Psychologists	1.113	.326	.953	.280	1.111	.237
Open Housing Laws	2.027	.118	2.806	.411	2.512	.577
Pollution	.917	.740	.582	.782	.805	.872
Strikes	2.058	.497	2.098	.484	2.464	.880
Computers	1.657	.298	1.570	.400	1.357	.914
Law and Order	1.929	.179	2.218	.138	1.688	.303
Interracial Marriage	1.775	.428	2.355	.456	2.468	.400
Disruptive Protests	2.354	.224	2.619	.394	2.452	.528
Small Businessmen	1.126	.095	1.301	.240	1.331	.161
Wire Tapping	2.512	.134	2.512	.460	2.822	.842
Labor Unions	1.501	.179	1.792	.232	1.819	.407
Large Corporations	2.030	.228	2.131	.643	1.774	.385

Table 13.--Continued.

	Condition					
	Fixed adjective		Fixed concept multiple dimension		Fixed concept evaluative only	
	$\overline{\sigma^2_X}$	$\sigma_{\sigma^2_X}$	$\overline{\sigma^2_X}$	$\sigma_{\sigma^2_X}$	$\overline{\sigma^2_X}$	$\sigma_{\sigma^2_X}$
Integration by school bussing	2.808	.401	3.097	.432	3.401	.423
Police	1.850	.354	2.274	.303	1.710	.391
Boycotts	1.233	.141	2.030	.363	2.087	.374
Civil Rights Movement	1.106	.173	1.458	.251	1.978	.535
Hippies	2.012	.392	1.643	.268	1.824	.212
Use of sit-ins, lie-ins, etc. by Civil Rights Demonstrators	1.373	.392	2.518	.396	2.496	.443

Table 14.--Means and standard deviations of $r_{X_1X_2}$ within each concept, across the ten interevaluative adjective correlations.

	Condition					
	Fixed adjective		Fixed concept multiple dimension		Fixed concept evaluative only	
	$\bar{r}_{X_1X_2}$	$\sigma_{r_{X_1X_2}}$	$\bar{r}_{X_1X_2}$	$\sigma_{r_{X_1X_2}}$	$\bar{r}_{X_1X_2}$	$\sigma_{r_{X_1X_2}}$
Migrant Workers	.39	.18	.43	.09	.27	.06
Draft	.59	.13	.56	.11	.39	.18
Psychologists	.49	.12	.42	.12	.50	.13
Open Housing Laws	.69	.08	.66	.15	.55	.13
Pollution	.17	.16	.29	.16	.15	.14
Strikes	.67	.09	.62	.06	.39	.12
Computers	.52	.12	.47	.16	.40	.20
Law and Order	.59	.10	.60	.12	.55	.09
Interracial Marriage	.52	.13	.60	.17	.50	.17
Disruptive Protests	.68	.08	.64	.12	.52	.13
Small Businessmen	.43	.10	.49	.05	.56	.13
Wire Tapping	.76	.06	.60	.10	.51	.14
Labor Unions	.65	.08	.62	.09	.54	.13
Large Corporations	.52	.12	.59	.08	.43	.13

Table 14.--Continued.

	Condition					
	Fixed adjective		Fixed concept multiple dimension		Fixed concept evaluative only	
	$\bar{r}_{X_1X_2}$	$\sigma_{r_{X_1X_2}}$	$\bar{r}_{X_1X_2}$	$\sigma_{r_{X_1X_2}}$	$\bar{r}_{X_1X_2}$	$\sigma_{r_{X_1X_2}}$
Integration by school bussing	.76	.03	.75	.11	.68	.08
Police	.65	.07	.70	.08	.62	.06
Boycotts	.58	.09	.68	.08	.68	.04
Civil Rights Movement	.46	.16	.57	.09	.62	.09
Hippies	.68	.07	.61	.12	.54	.12
Use of sit-ins, lie-ins, etc. by Civil Rights Demonstrators	.51	.12	.64	.10	.55	.15

Table 15.--Obtained values of $\bar{\sigma}_X^2(1-\bar{r}_{X_1X_2})$ for each concept for the fixed adjective, fixed concept, multiple dimension and fixed concept, evaluative only conditions.

$\bar{\sigma}_X^2(1-\bar{r}_{XX})$	Condition		
	Fixed adjective	Fixed concept multiple dimension	Fixed concept evaluative only
Migrant Workers	1.133	.972	1.068
Draft	.937	1.075	1.714
Psychologists	.566	.548	.556
Open Housing Laws	.618	.962	1.120
Pollution	.764	.412	.688
Strikes	.675	.795	1.498
Computers	.789	.824	.816
Law and Order	.797	.880	.751
Interracial Marriage	.854	.932	1.222
Disruptive Protests	.744	.943	1.174
Small Businessmen	.642	.664	.587
Wire Tapping	.615	1.002	1.383
Labor Unions	.521	.675	.822
Large Corporations	.964	.875	1.015

Table 15.--Continued.

$\overline{\sigma_X^2}(1-\overline{r}_{XX})$	Condition		
	Fixed adjective	Fixed concept multiple dimension	Fixed concept evaluative only
Integration by school bussing	.674	.777	1.088
Police	.653	.691	.658
Boycotts	.517	.654	.676
Civil Rights Movement	.594	.621	.758
Hippies	.654	.641	.835
Use of sit-ins, lie-ins, etc. by Civil Rights Demonstrators	.673	.904	1.118

APPENDIX III

APPENDIX III

This appendix will derive the formula used to correct the regression lines for attenuation. The key to the derivation is the fact that the sampling error in estimating a population variance is analagous to the unreliability in estimating a subject's true score. That is, when one set of sample variances is plotted as a function of the other, the sampling error equation

$$\hat{\sigma}^2 = r^2 + e$$

is exactly analagous to the usual

$$X = T + e$$

The population variance in the true score and the sample variance differs from the population variance by sampling error. When you move from concept to concept, the sampling errors are independent of one another in the same sense that error scores for different subjects are independent. Furthermore, if the population variances for two different groups satisfied a linear equation

$$\sigma_2^2 = \alpha\ \sigma_1^2 + \beta$$

the correlation between sample variances would be attenuated by sampling error by exactly the same amount as the traditional

$$r_{XY} = \sqrt{r_{XX}\, r_{YY}} \quad \text{if} \quad r_{T_X T_Y} = 1$$

If we were correcting the regression line for gwo variables for attenuation, then the necessary equations would be

$$E\,(T_x) = E\,(x) \approx \bar{X}$$

$$E\,(T_y) = E\,(y) \approx \bar{Y}$$

$$\sigma^2_{Tx} = \sigma^2_x \cdot r_{xx} \approx s^2_x \cdot r_{xx}$$

$$\sigma^2_{Ty} = \sigma^2_y \cdot r_{yy} \approx s^2_y \cdot r_{yy}$$

$$r_{T_x T_y} = \frac{r_{xy}}{\sqrt{r_{xx}}\,\sqrt{r_{yy}}} \approx \frac{r_{xy}}{\sqrt{r_{xx}}\,\sqrt{r_{yy}}}$$

The five parameters for true score would then be used to generate a regression line for true scores in the usual fashion.

Of the numbers on the left side of these equations, $\bar{x}$, $\bar{y}$, s^2_x, s^2_y, and r_{xy} would be the usual statistics calculated for the two sets of sample variances as if they were two sets of scores. What about r_{xx}? The first step in the derivation is to shift from the usual

$$r_{xx} = \frac{\sigma^2_T}{\sigma^2_x}$$

which requires σ^2_T, to the formula

$$r_{xx} = \frac{\sigma^2_x - \sigma^2_e}{\sigma^2_x}$$

which requires σ_e^2. The σ_e^2 in this equation is the variance of sampling error and will be estimated using traditional statistical formulas. Thus for a single concept,

$$\sigma_e^2 = \frac{\sigma^4 (k-1)}{n - 1}$$

where n is the sample size, σ^2 is the true variance for that concept and k is the kurtosis for that concept. If the distribution for that concept was normal, this would produce the usual

$$\sigma_e^2 = \frac{2 \sigma^4}{n - 1}$$

Since e has an (unobserved) true mean of zero for each concept, the variance of e over concepts is the mean of its variance within concepts, i.e.

$$\sigma_e^2 = \frac{2 \overline{\sigma^4}}{n - 1} = \frac{2 (\overline{(\sigma^2)^2})}{n - 1}$$

Using $\bar{x}$, σ_x^2, etc. for the scatterplot statistics, this means that

$$\sigma_e^2 = \frac{2 \overline{x^2}}{n-1} \approx \frac{2(\bar{x}^2 + \sigma_x^2)}{n - 1}$$

and hence

$$r_{xx} = \frac{\sigma_x^2 - \sigma_e^2}{\sigma_x^2} \approx \frac{\sigma_x^2 - \frac{2\bar{x}^2 + 2\sigma_x^2}{n - 1}}{\sigma_x^2}$$

$$\approx \frac{(n-1)\ \sigma_x^2 - 2\bar{x}^2}{(n-1)\ \sigma_x^2} - \frac{2}{n-1}$$

$$\approx \frac{(n-1)\ \sigma_x^2 - 2\bar{x}^2}{(n-1)\ \sigma_x^2}$$

This last formula is the one used on the body of the text.