ADVERBS MULTIPLY ADJECTIVES

Norman Cliff *University of Southern California*

Is IT possible that the difference in the meanings of the adjectives *nice* and *unpleasant* is a simple numerical one? If so, what happens when the words are modified by an adverb: *very nice, very unpleasant, somewhat nice, somewhat unpleasant?* Are the meanings of the combinations changed by arbitrary amounts or are the changes systematically related to the meanings of the unmodified words? For a number of years, but especially since the early fifties, there have been studies of questions like this by statistically minded psychologists and linguists. In this essay, we describe one such study and its results. It showed that "adverbs multiply adjectives" in a very literal sense.

There does seem to be an analogy between adverbial modification and multiplication. Compare the meanings of *very nice* and *very unpleasant* to *nice* and *unpleasant* by themselves. The modified pair are more different from each other than the individual adjectives. This could be explained by assuming the *nice* represents a positive number (+2, for example), *un-*

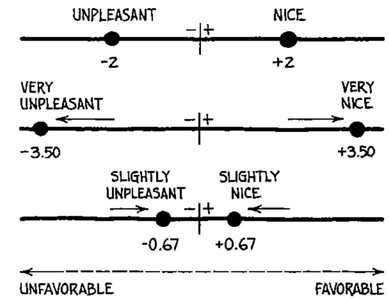

FIGURE 1
The upper part is the hypothetical position of two adjectives on a continuum of favorableness-unfavorableness when they are not modified. The middle portion shows the effect of modification by very *and the lower the effect of modification by* slightly

pleasant a negative one (say, —2), and *very* a positive number greater than one (say, +1.75); then the meaning of the combination is the product of the adverb and adjective numbers. (For example, the meaning of *very nice* would be $1.75 \times 2 = 3.50$.) The effect is illustrated in Figure 1, where the top portion represents the hypothetical positions of two adjectives on a continuum when unmodified, and the middle their positions when modified by *very*. The effect of adverbs such as *slightly* and *somewhat* that reduce the intensity of the adjectives they modify can be explained by assuming that the numbers they represent are less than unity, as illustrated in the bottom portion of Figure 1, where we assume *slightly* has a value of 0.33, so that *slightly unpleasant* has a value of $0.33 \times (-2) = -0.67$. Various adjectives occupy different positions on the continuum, and various adverbs have different "multiplying values," so that the function of the adverb is to move adjectives up or down the continuum in a regular way.

Our idea is translatable into a simple mathematical formula:

$$f = ia.$$

It states that the favorableness f of a combination is the product of the intensifying effect i of the adverb times the number a representing the adjective.

QUANTIFYING FAVORABLENESS

The first problem in testing such a theory is quantifying favorableness. Here, as in other similar studies, we started with a fairly straightforward approach based on the judgment of native speakers of the language, but judgments gathered in a formalized way. We defined a scale of favorableness, and 11 categories on this scale. Words and combinations of words were given to a set of judges, who in this case were college students, and each judge was instructed to indicate the category which seemed appropriate to him for each word. The categories were described as running from "most favorable" through "neutral" to "most unfavorable." The setup is illustrated in Figure

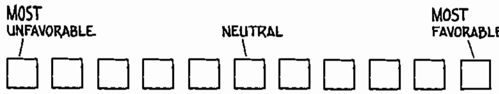

FIGURE 2

Format for judging the combinations. The pattern was repeated down each page, with adjectives or adverb-adjective combinations listed at the left and the boxes printed on the right. The headings "most unfavorable," etc. were given only at the top of the page

2. Then, after the data had been gathered, the numbers —5, —4, . . . , 0, . . . , +5 were assigned to the categories, and the average of the numbers assigned by the judges was calculated.[1] We found that there was good agreement between judges on the rating for each word or combination of words.

In the study of adverb-adjective combinations there were 150 words and combinations of words judged in this way. These were all combinations of the 9 adverbs in Table 1 with the 15 adjectives in the table, plus the adjectives alone. Average ratings of the kind described in the previous paragraph were calculated for all these 150 words and combinations.

DERIVING ADJECTIVE AND ADVERB NUMBERS

The 150 average ratings from the study could be used to test the multiplicative formulation and to derive the adverb and adjective numbers. In the case of the ratings of the adjectives by themselves, we assume that $f = a$ since there is no adverb involved. In order for the idea of multiplicative combination to have much use, we need to assume further that the number which the adverb represents does not change from adjective to adjective, and vice versa, so that we should be able to find a number for each adverb and one for each adjective.

People are, of course, not entirely consistent either with themselves or with other people in the way they rate the adjectives; thus no such mathematical formulation can be expected to fit the data exactly, so we need to have procedures for the statistical estimation of the adverb and adjective numbers that are most consistent with the data. We also need statistical means of seeing the degree to which the formulation is consistent with the data: is it way off the mark, a rough approximation, or a very accurate description?

The method of estimating the numbers revolved around comparing the favorableness of unmodified adjectives to their favorablenesses when modified by a particular adverb. In Table 2, we have the favorablenesses of five adjectives when by themselves and when modified by *decidedly*. In theory, all we have to do is divide the favorableness of, for example, *decidedly admirable* by the favorableness of *admirable* to get the intensifying value of

[1] In the initial study of this kind, the statistical analysis that was used was much more elaborate than simply taking averages and required several weeks of computation using the equipment then available. Subsequent events showed that virtually the same results would have been achieved by using the simple averaging process. For the sake of simplicity and because the results would be equivalent, the study is described as if the simple averaging process had been used.

TABLE 1. Adverbs and Adjectives Used in the Adverb-Adjective Combinations

ADVERBS	ADJECTIVES
Slightly	Evil
Somewhat	Wicked
Rather	Contemptible
Pretty	Immoral
Quite	Disgusting
Decidedly	Bad
Unusually	Inferior
Very	Ordinary
Extremely	Average
	Nice
	Good
	Pleasant
	Charming
	Admirable
	Lovable

TABLE 2. Average Ratings of Five Adjectives

ADJECTIVE	UNMODIFIED (a)	MODIFIED BY DECIDEDLY (ia)	ESTIMATE OF INTENSIFYING VALUE (i)
Disgusting	−3.10	−3.42	1.10
Inferior	−1.94	−2.70	1.39
Ordinary	−0.35	−0.63	1.80
Pleasant	2.18	2.75	1.26
Admirable	2.92	3.33	1.14

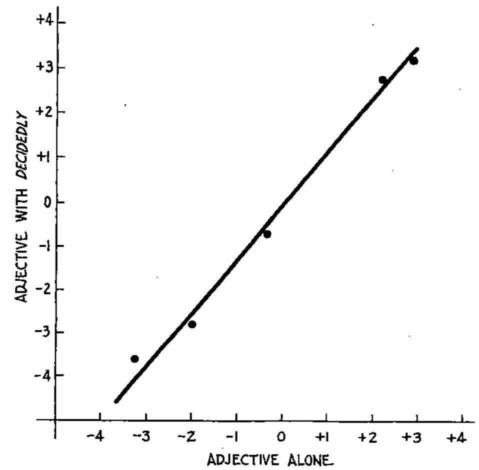

FIGURE 3

Plot of the data in Table 2 in which the mean judgment of each adjective used alone is the horizontal coordinate of each point and its vertical coordinate is the rating of the same adjective modified by decidedly. *Thus, the lowest left point is for* disgusting *by itself (−3.10) and* decidedly disgusting *(−3.42)*

decidedly because this would be ia/a; the a's would cancel out, leaving us with an estimate of i for *decidedly.* In this case, that would be 1.14. The trouble with this approach is that we get different values for i depending on which adjective we pick to work with, as we see in the last column of Table 2.

Statistical methodology gives us a way out of this inconsistency. It provides a way of arriving at an overall best estimate of the i for any adverb. To illustrate the way it does this, the data for the five adjectives are plotted in Figure 3, where the unmodified rating (horizontal axis) of each is plotted against the rating when modified by *decidedly* (vertical axis).

Another way to state the multiplicative principle is to say that the favorableness when modified should be proportional to the favorableness when unmodified, with the intensifying effect of the adverb as the constant of proportionality. This means that the points in the figure should lie along a straight line through the origin, and that the numerical value of its steepness (the tangent of the angle it makes with the x axis) is the adverb number. It can be seen that the points in the figure lie pretty close to a straight line. The statistical method called *least squares* allowed us to define the line that is *closest to all the points,* and we used the slope of this line as the most representative value of the intensifying effect of *decidedly.*

This is an example typical of a number of applications of statistical methodology. Whenever, as here, we are trying to fit a model to some observations, there are discrepancies between the model and the data. The model usually states the general form of a relationship but leaves one or more constants (or parameters) to be determined from the data. In this case, our model tells us that the relation should be linear and through the origin, but it doesn't

TABLE 3. Intensifying Values of Adverbs and Scale Values of Adjectives

ADVERB	i	ADJECTIVE	a
Slightly	0.54	Evil	−2.64
Somewhat	0.66	Wicked	−2.54
Rather	0.84	Contemptible	−2.20
Pretty	0.88	Immoral	−2.48
Quite	1.05	Disgusting	−2.14
Decidedly	1.16	Bad	−2.59
Unusually	1.28	Inferior	−2.46
Very	1.25	Ordinary	−0.67
Extremely	1.45	Average	−0.79
		Nice	2.62
		Good	3.09
		Pleasant	2.80
		Charming	2.39
		Admirable	3.12
		Lovable	2.43

tell what the slope of the line should be. Among all the lines that could be used, it is natural to pick the one that gives the least discrepancy between the model and the data; geometrically, that means drawing the line that is the closest to the points. There are various ways of mathematically defining "closest," and to say we used least squares means we used a particular one of these. In general, it means picking the value of the parameter that gives the smallest total *squared* deviations between the model and the data. Here, that means finding the line for which the sum of the squared distances from the points to the line is the smallest.[1] Just how this is done depends on the exact nature of the application, and, while the basic aspects of the method require only elementary calculus, there are a number of mathematical and computing tricks that have been developed. As applied here, it allowed us to find "best possible" estimates of the multiplying value of the adverbs.

The process was repeated using each adverb; the slope of the line that most nearly related the favorableness of adjectives combined with it to their favorableness when used alone was determined. These numbers are the adverbial multiplying values presented in Table 3. The numbers indicate that *slightly* and *somewhat* diminish the effect of an adjective to about half its original value; *quite* has almost no effect at all (in contrast to its dictionary

[1] To a large extent, the reason for using the sum of squared discrepancies rather than some other criterion such as the sum of absolute discrepancies has to do with mathematical convenience.

meaning); *extremely* has the greatest intensifying effect, making an adjective about one and a half times as strong as it would be alone.

We could use the average ratings of the unmodified adjectives as their *a* values. These are not the best estimates, however, since they place too much reliance on a single set of data, whereas, according to our formula, the number represented by an adjective enters into the ratings of *all* its combinations. Therefore, once we have determined the adverb numbers we can use them in a statistical procedure analogous to that used in fitting the straight lines to arrive at optimum or best-fitting estimates of the adjective numbers. These are the ones given in Table 3, where we see that the unfavorable adjectives represent negative numbers, and the favorable ones positive numbers. Two adjectives, *ordinary* and *average,* have numbers near 0. Perhaps this is why it seems odd or funny-sounding when they are modified by one of our adverbs: zero multiplied by anything is still zero, so why bother?

CHECKING THE FIT OF THE MODEL

There were statistical checks on the accuracy of our formulation in all of the steps performed to this point (the closeness of points to the lines, for example), but one final, overall evaluation was made. This was to multiply each adverb number by each adjective number and compare the result to the average rating for that combination. [For example, the rating of *very contemptible* should be $(1.25) \times (-2.20) = -2.75$.] This tells us the accuracy with which the 24 derived numbers (9 adverbs plus 15 adjectives) could be used in conjunction with our multiplicative theory to account for the 150 data numbers. When this was done, it was found that the average (unsigned) discrepancy was about 0.15 of a category, so that a combination with actual average rating of say 3.00 might well come out as about 3.15 or 2.85 when we multiply its adjective number by its adverb number. In some cases it would come out closer, and in some cases somewhat farther away, but on the average the unsigned discrepancy would be about 0.15. Since the data we started with were the average categories in which word combinations were placed—a task that many people might think cannot be done consistently with single words much less with combinations—and since the theory tested is that the use of words in communication is really a numerical process, the closeness with which the data fit the theory was surprising.

One additional point may be worth noting. This has to do with the assignment of the numbers —5, —4, etc. to the categories. This might seem very reasonable intuitively, but it could be that the lowest category rather than the middle one should be zero and the categories should be numbered 0, 1, 2, up to 10 if we are to interpret the labels as favorableness numbers. This would be analogous to having used degrees Centigrade when we should be using degrees Absolute in measuring temperature in a gas-law experiment.

The fact is that an aspect of the statistical analysis validated our locating the zero point in the middle category rather than at the end of the scale or some other place. In effect, we were able to use the analysis to locate the *absolute zero point,* and it turned out to be in the middle category.

REPEATING THE STUDY

The same experiment was repeated with two more groups of judges in other parts of the country with an equivalent degree of verification of the principle of multiplication. There was some variation from group to group in the actual adverb and adjective numbers obtained, presumably as a reflection of slightly different local usages. The study was further repeated in Australia and Great Britain, again with very similar results. Since then, the word lists have been translated into French, Polish, Spanish, Norwegian, German, Japanese, and other languages, and in all cases the principle that adverbs multiply adjectives has been verified, although the words' numerical values sometimes surprise the translators. The same sort of principle has also been shown to apply to other classes of words such as adverbs denoting probability, and the tenses of the verb *to be*. Two adjectives, on the other hand, combine in a way that is more like averaging than multiplication. Each adjective retains its own characteristic effect when used in combination with another regardless of what it is combined with, but two adjectives together do not have anywhere near the product or the sum of the effects of each separately.

The idea that words combine numerically, that they even multiply together, is now well established, but it originally seemed rather farfetched. To find that people agree fairly well on the numerical value to assign to a particular word was in itself surprising the first time or two it was done. To find that they did the same thing with combinations of words was an additional surprise. To find that the numbers they assign to the combinations conform to the multiplicative rule initially seemed outlandish. In fact, data that could have been used to establish the multiplicative rule was published some 15 years previous to the study described here, but no one thought of looking at it that way, probably in part because the relevant statistical methods were not widely enough known.

The final conclusion reached in the studies described here was established as the end result of a large amount of statistical methodology employed at numerous stages in the process. It started with the tabulation of the original responses by the judges, and ended with the overall estimate of how well the multiplicative hypothesis agreed with the data. It may be worth noting in passing that the process used here is completely analogous to verification of the laws interrelating the temperature, pressure, and volume of gases, including a check on the validity of the numbers assigned to the categories (analogous to estimating absolute zero on the temperature scale). The inexact nature of the data used here makes statistical methodology play a larger part in this study than it does in most physical science problems, and the use of statistical methods was essential in this work.

REFERENCES

The study described here is also reported informally in N. Cliff. 1958. "Intensive Adverbs from a Quantitative Point of View." *College Composition and Communication* 9:20–22. A more technical report is "Adverbs as Multipliers." 1959. *Psychological Review* 66:27–44.