International Journal of Orofacial Myology and Myofunctional Therapy
Official Journal of the International Association of Orofacial Myology
50th ANNIVERSARY

Volume 2 | Number 3 | pp. 54-66 1976

Tutorial

Analysis and interpretation of research publications

James M. Wood

Suggested Citation
Wood, J. M. (1976). Analysis and interpretation of research publications. *International Journal of Oral Myology, 2(3)*, 54-66.
DOI: https://doi.org/10.52010/ijom.1976.2.1.8

The International Journal of ORAL MYOLOGY
Volume 2, Number 3 **July, 1976**

Analysis and Interpretation of Research Publications

James M. Wood, Ph.D.
Omaha, Nebraska

Dr. Wood is on the faculty of the University of Nebraska at Omaha

The word "research" is not a new one to any of us. We see it used on television, in the newspapers, on the radio, in magazines, in grocery stores, on billboards, in journals and over the backyard fence. It is used as an advertising gimick to sell low tar and nicotine cigarettes, cars with low gas mileage, detergent with high wash mileage, cereal with high protein content, margarine with low cholesterol, and mouthwash which kills germs by millions.

We are touched by research in almost everything we do. Yet, most of us know little about the research which has been done or if it has any real merit. This is no less true of the field of oral myology. Perhaps it is *especially* true of this new field of study.

If you think about the research to which you are exposed in a typical day you may realize that it takes different forms depending upon what the researcher is studying. The comparison of the smoothness of the ride one gets in a Ford Granada versus that in a Mercedes Benz is quite different than the opinion polls obtained by George Gallup.

Perhaps the first point which should be made to you is to encourage you to discover that what you are reading in the field of Oral Myology, Dentistry, Speech Pathology or some other field is a report of *research* and not someone's opinion *about* research or about a topic generally. Research implies a scientific investigation based upon certain established principles. It is an orderly persuit of knowledge. This is not to say that opinion papers or summaries of research are worthless. But you should keep in mind that when you read these papers you are also reading something which may be biased by the writer's own needs, philosophies or attitudes about the topic.

Research studies fall into one of three broad categories—historical, descriptive and experimental. Each of these forms of research will be discussed so that you know what techniques are used in each and which types are applicable to the topic under investigation. I should say at the outset that no method is superior to another—just different. Where one type of research is appropriate another is inappropriate. Its similar to comparing horses to apples. Each has a different function and therefore no meaningful comparison may be made. One could argue that you could study how many apples a horse eats in a day, but this doesn't compare horses and apples. They're different, and different methods must be used to study them.

Historical Research

Historical research is often thought of as the poor step-child of the other two. After all, it is merely the rehashing of old ideas. No new information is gained from historical research—right? Wrong. What historical research attempts to do is put into order prior information in a new and meaningful way. Its similar to a cook who, faced with a variety of left-overs, discovers a new and exciting dish. Realistically this happens as rarely to the cook as it does for the historical researcher. But when it does happen the excitement is superb. Can you imagine the thrill of a student of ancient maps and sea-going vessels discovering the most plausible location of a sunken ship loaded with treasure? Romantic? Yes, but it does happen doesn't it?

In the field of Oral Myology historical research is almost nonexistent. The reason is obvious—the field is too new. It is now as new as many of us suppose, however, when we realize that it began in the United States with the writings of Dr. Rogers about 1917. So we do have an historical legacy.

It may be said that a review of the literature is historical research. It is, but usually it is specifically oriented to a particular topic to be investigated by another method—either descriptive or experimental. The perspective is usually somewhat different. In most historical research the evolution of the problem is studied along with probable cause and effect relationships of that evolution. Most literature reviews look only at the relevance of a topic to the one under investigation.

As you all know, the road by which oral myologists have traveled in their attempt to help their patients has not been a smooth one. We are all familiar with the criticisms about the work we do. It may have been helpful if someone had done an historical study of the rise and fall of oral myofunctional therapy when Rogers began to write about it. If you examine the literature you will see a paucity of information from Rogers' first article and the end of World War II. Why? This is not for me to answer here because I don't know. It may have been due to the depression era. It may have been due to the newness of the field of orthodontics—one that is not old, by-the-way, as speech pathology. Whatever the cause, some insights into that early history may have better equipped us to cope with the recent controversial developments in this field.

Descriptive Research

Descriptive research, the next method to be discussed, is highly familiar to us all. Most marketing research is of this type. This tells the manufacturer who buys the product, how they heard about it, where they bought it, how old the buyers were, the sex of the buyers, their income bracket, the size of the city they live in, what the occasion was for buying it, etc. In this way the manufacturer knows where, how, when, and to whom to advertise his product. Descriptive research deals with the here and now. It does not deal with the past nor does it attempt to alter the future. It analyzes things as they are at the present time.

Descriptive research uses two common techniques to collect data. One is a survey or questionnaire method. This is the most common type and one you see used all the time around your home. The little cards a manufacturer has you complete to validate the warranty is often a small questionnaire. The call you get asking if you are watching a certain TV program uses the survey technique. Census takers use questionnaire to gather information about a city's population.

Surveys using questionnaires seldom include an entire population of people from which to gather data. This would be much too time consuming and costly. The census is one exception to this rule. Therefore, since in most cases the entire population cannot be studied, the researcher must decide how to select a smaller group which is most representative of that larger group or the entire population. He knows he runs the risk of not studying a representative sample of the population, but if he is careful about how he selects the sample and makes sure it is large enough, his chances of making an error are small.

I believe it is best if I now present some terms which I will be using to explain both descriptive and experimental research. I have already used some of them. Here they are again plus a few more which you will see throughout the remainder of this article:

Population—	A group of people having a common characteristic which the researcher wishes to study.
Sample—	A selection of members of a population used as a convenience to the researcher because it is smaller and more easily studied. to the researcher because it is smaller and more easily studied.
Correlation—	The degree to which data are similar when compared to one another.
Normal distribution—	The way in which most data is arranged in a population. Typically this is represented by a bell shaped curve (See Figure 1.), where the mean, median and mode or measures of central tendancy are identical.
Mean—	An average where a numerical unit or score for each member of a sample is divided by the total members (N) in the sample.
Median—	The best representative number or score the middle of a sample when those numbers or scores are ranked from the highest to the lowest.
Mode—	This is the most common recurring number or score in a sample.
Distribution—	The array of numbers or scores which represent the members of a sample.

Before we get too far, let's take a look at how we may apply some of these terms. Suppose we set up a small survey to see if children of age five suck their thumbs while asleep at night. We decide to check a hundred kids. To make sure we do the job properly we want to select children which typify the community in which they live. If we are studying children in a large city we may wish to control such factors as socioeconomic status and race or cultural background. So we may choose to select our sample from various kindergarten classrooms around the city. Such factors are called "variables" since they are circumstances which may influence the results of the study and must either be accounted for or controlled. We will speak more specifically about variables later on. Another variable we should consider is the sex of the children selected. Choosing too many boys or too many girls may bias our data and render it invalid. So in our study we will choose fifty boys and fifty girls. Now so far we have controlled the variable of sex ratio and have taken into account the socioeconomic and cultural background of the children. One variable we have not considered is the sibling order for the families from which our children are selected. Here is where we may consider random sampling. Random sampling is a technique used to insure that a sample is not chosen with certain "build-in" biases. In our study we have not chosen a totally free random sample. We have made sure that each socioeconomic level and cultural background is represented. We have also made sure that we have an equal number of boys and girls in the sample. Had we chosen to do so, we could have obtained a list of all the five year old children in the city and selected our sample from this list without regard to any of the variables we have discussed previously. In this instance we would not have controlled for the variables of sex, race or cultural background or socioeconomic status. Therefore, when we analyzed our data we could not have said anything about these variables unless our questionnaire covered them. In the case of sibling order we have chosen not to control this variable. We can't say that we have "randomly chosen" our sample for this variable, but the effect on our data will be similar to it. In other words, we will expect that with one hundred children we will probably get a rather normal distribution curve—a bell shaped curve—with regard to the variable of sibling order (see figure 1). I believe I should explain a point here. The choice of a random sample is used for an entire study in most cases. That is, you don't choose to control several variables and then decide to select a random for the rest. If we had chosen to do a random sampling in our study we could have sampled randomly from all the five year old children in our city. But this would not have accounted for a sex bias in our sampling since it would be possible, although not probable, that the entire sample would be made up of boys (or girls). It could have selected all the children from a high socioeconomic area (or a low one). In other words, unless the sample is quite large, a researcher is better off controlling the variables he feels may bias his data rather than leave such possibilities to chance.

Now that we have selected our sample, what will we do with it? Well, since we want to know which kids suck their thumbs at night and which ones don't, we need some way to observe that behavior. Since we can't go into each child's house and watch him, we will ask the parents to watch him for us. We will give each parent a chart to complete each night. It will be simple to fill out (see figure 2). The parents will be asked to observe the child three times each night for two weeks. They will be told that sucking should be checked if the child has him thumb in his mouth at all whether they see movement of the mouth or not. This will be our operational definition of thumb sucking for this study. Note that we have not included finger sucking in our study. These observations will be made just after the child has fallen asleep, at least one hour after that, and once again before the

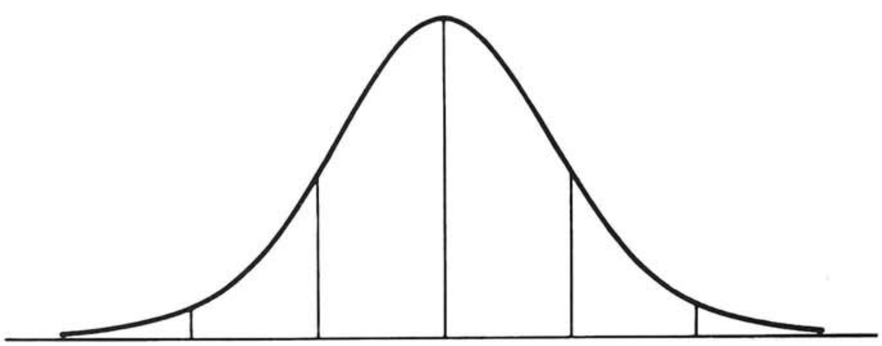

Fig. 1. A normal distribution or "bell shaped" curve

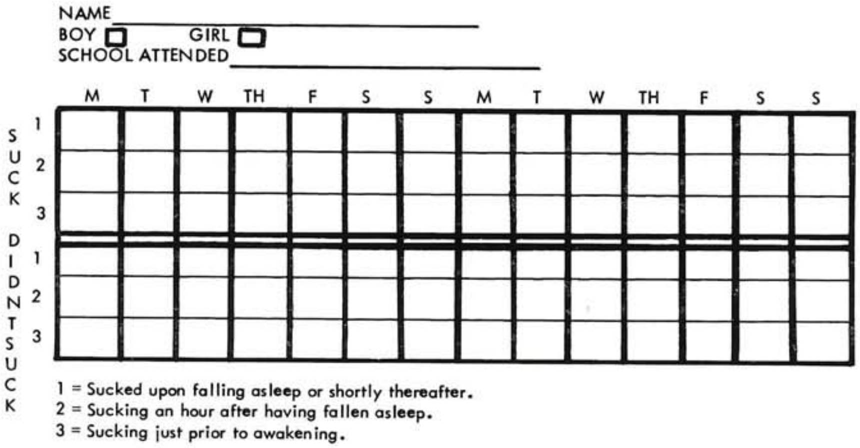

Fig. 2. Chart used by parents to tally thumb sucking behavior of the child

child awakens in the morning. This will enable us to make sure the child has or has not sucked, when and for how long. Two weeks of observation will give us some confidence that we are seeing the typical sucking behavior of a child. Now you see, we are making some assumptions about sucking and about what constitutes typical behavior. Many disagreements occur from such assumptions. But remember an important principle. Good research reporting will state the assumptions somewhere. The danger, of course, is that the researcher will assume that an assumption is so obvious or logical that it doesn't need to be stated and the reader will therefore "know" what was in the researcher's mind. There is another danger, of course. That it that the assumption is correct when it is not.

We will assume for a moment that the data has been collected and is ready to be analyzed. What do we do with it. In this case the analysis is quite simple. We may compare the number of children who suck with those who don't. Then, if our data gives us enough children who suck their thumbs, we may be able to say something about these children. Several questions may be posed: Does socioeconomic status affect sucking habits? Does race/cultural background affect sucking habits? Do boys or girls suck the most? Does sucking occur most frequently during the first part of the night? After the child has "settled" in his sleep pattern? Just prior to awakening? The questions may be addressed with the use of averages or means for these children. Anyone may ask "does your study answer these questions?" The answer is "no", at least not for the entire population of five year old children in our city. But it does answer the questions insofar as our sample is concerned. In addition, it does point to a trend which bears watching. If our data were to show, for example, that children who suck their thumbs at night do so most often upon falling asleep, we may feel that a therapeutic strategy to alleviate the habit must concentrate its energies on the child at that period of time.

I should say something here about research in general. Seldom in human behavioral research is the data so overwhelming that the one and only "answer" is discovered. This is particularly true for surveys.

Surveys are not the only type of descriptive research. Another popular technique is to compare two procedures to see if they are doing the same thing in two or more different ways. In other words we study their relationships. Correlation studies are of this type. Let's get away from oral myology for a moment and let me share with you a little study we have just completed at the University of Nebraska at Omaha Speech and Hearing Clinic. In our routine evaluations of language impaired children we have given a picture vocabulary test for years. We have used this as an index of a child's ability to handle incoming language or that which he hears and sees as opposed to that which he speaks or writes. Last year we added another test to the battery which assesses not only receptive vocabulary but also the structure of language (i.e. the comprehension of nouns, verbs, adjectives, adverbs, conjunctions, etc. and their relationships in a sentence). We felt that although the two tests were assessing "receptive" skills the scores were different and thus must be measuring "different" receptive skills. In an effort to test this feeling we designed a very simple correlational research study. We decided to have all children who came to our clinic who were of the age to be tested with these tests receive both of the tests. Since both tests were receptive tests and were also verbal tests we wanted to see how they correlated. We were surprised at the outcome. Our observations that we were seeing gross differences in the scores of the two tests on the same child and thus the tests must not be correlated was wrong. The correlation coefficients we obtained were very high—upwards of .90. The significance level of these correlation coefficients was between .01 and .001 for these correlations.

I have just introduced two new terms to you—correlation coefficients and significance levels. We should now look at these two terms because they are used frequently in correlation studies and in experimental research. Remember that a correlation is the degree to which data are similar when compared to one another. A correlation coefficient is a way to express this relationship numerically. It is arrived at by way of statistical formula. High correlations represent a close relationship of the data being studied while low ones represent a poor relationship. There is another factor which may confuse you. That is, it is possible to have a negative correlation. In the case of our study at UN-O had we found any negative correlations we may have assumed that a particularly low score on one of the tests would point to a possibility of a very high score on the other one. This happens rarely in behavioral research since most human behaviors are related to one another, at least to some degree. I don't know if anyone has studied this problem, but it is possible that the physician's skill as a physician and his penmanship are inversely correlated!

Correlation coefficients of .85 are considered significant for individual predictions of correlation. Correlation coefficients of .60 are considered significant for group predictions of correlation. The reason for this difference is that there is a greater probability for variance in a group than in an individual. Therefore it is permissable to accept a lower correlation coefficient.

Significance levels are determined on the basis of the relationship we have determined having occurred by chance. Now we get into the term "probability". In other words, if we did this study over and over again, what is the chance that we would find results which were quite different. If, as we find in behavioral research, the .05 level is used, we may say that there are five chances in a hundred that we would get different results than we got this time. Put another way, if the same study was done 100 times we would get the same results 95 times. If the .01 level is used the chances of getting different results goes down to only one in a hundred. If the .001 level is used, the chances for getting different results are one in a thousand. Most behavioral researchers will accept the .05 level. The acceptance of a level is determined by the researcher. Therefore, when I told you of the results of our study at UN-O having significance levels of .01 to .001 you may see that the two tests were highly related.

Descriptive research also is used to establish norms. The tables a dentist used to know when certain teeth will erupt in a childs dental arch are products of this type of research.

Experimental Research

Experimental research is the third method I will present. I have saved it for last because it is perhaps the more complex of the three. Whereas descriptive research deals with the "here and now," experimental research attempts to change something and then analyze that change. Again we deal with probability. We want to know to what extent the changes we observe may have occurred by chance. Before I go further into this topic I believe I should present a basic principle of scientific inquiry. It is assumed that scientists are honest people. This may or may not be the case, but lets assume for a moment that its true. Being realistic we also know that since scientists are human beings they are subject to personal biases as are we all. Therefore, scientific investigation has set for itself an assumption about all that it does to change something. It assumes that no change will occur—that as a result of the research, no change will have been made. Now this attitude seems ridiculous, doesn't it? Why would any researcher set up an elaborate study to effect a change of some sort and then systematically go about trying to prove he is wrong? The reason for this attitude is to attempt to eliminate bias. It helps the researcher maintain his objectivity with regard to the data he is collecting. If he has a psychological committment as to how it will all come out, he may inadvertently contaminate his data by his biases. Therefore, as a precaution against this possibility he states a "null hypothesis." The null hypothesis is a hypothesis of no difference in the data being analyzed. Basically the researcher says to himself, "after all this study when we analyze what we have done we will find that it has made no difference." Now this is not the exact way in which null hypotheses are stated, but it will serve our purpose here. In other words he has no committment as to what the results will be. "Let the chips fall where they may" so to speak.

There are situations where a researcher wishes to say that there will be a difference between one set of data and another. He may even want to say how the difference will occur. When this situation arises the researcher states an experimental hypothesis. In a sense, he is "going out on a limb." Because of this confidence in the outcome of the experiment, the researcher may set a higher significance level as being acceptable. If he had accepted the .05 level for a null hypothesis he may accept the .01 level for the experimental hypothesis. The reason for this decision lies in the statistical analysis of the data. The statistical differences in a study must be greater when a null hypothesis is tested than when an experimental hypothesis is tested. The reason for this is that if the researcher has a strong notion that a difference will be found he has, indeed, "stuck his neck out" and therefore is entitled to accept a smaller obtained difference in the data. There is a danger in this, however, and it is that since the researcher has a notion of the outcome of his experiment, he runs the risk that the difference he obtains doesn't really make any difference. Again, lets set up a small research project using an experimental design this time. I believe this will help you to understand the principles used in this type of research.

In this study we wish to determine if myofunctional therapy will correct or reduce class I open bite dental malocclusions in children ten years of age. Much as we did with the thumb sucking study we need to select a sample. The same considerations may need to be applied to this sample. We have some variables to control. Lets see if we can list them.

age
sex
class I open bite dental malocclusion
Tongue thrusters

There are other variables which may affect the outcome of our study. I am sure than if you think for a moment you may add them to the list. Since we are dealing with a learning task it would be well to make sure that the children are capable of the task to be learned. Intelligence should probably be considered. We may also need to consider if there is someone in the home to help the child with his exercises. Since we are dealing with the oral cavity and its related structures we may also wish to account for the space within the mouth and throat, the size and strength of the muscles involved as well as their coordination (i.e. the tongue, lips, cheeks and muscles of the jaws), the presence of any habit which may have an effect upon dental alignment (i.e. thumb or finger sucking, lip biting, etc.), the ability of a child to breathe through his nose, his and his parents cooperativeness and on and on.

As we contemplate these variables we realize that we may miss something which is crutial to the success of our study. How do we account for these variables? In most research the size of our sample helps us out. As a rule of thumb, as the size of the sample increases, the probability of an obtained difference occuring by chance decreases. In other words, the more children we have in our study the more reliable our results become.

Notice that in the previous few statements I have used the term "difference". What difference am I talking about. This involves the way in which the study is designed. Usually an experimental study of this kind uses two groups. The groups are identical for all intents and purposes or at least are distributed normally as far as the bell shaped curve is concerned relative to the variables which we are not controlling directly. Therefore, if any difference between the groups is found at the end of the study it must be due to something we did, because at the outset there were no differences. The three most common ways are: matched groups, random groups and accepted groups. A study may combine the methods. Matched groups use subjects who are essentially identical with regard to the variables we wish to control. In random groups, the assignment of subjects to one group or another uses chance probability. In other words, each subject has as great a chance being selected for one group as he has for another group. In accepted groups all available subjects are used. There are advantages and shortcomings to each method of group selection. There are many ways to set up these groups.

In our study we will choose two groups which are essentially identical relative to the four variables we listed first. That is, they must be children of the same age and sex and must have class I open bite dental malocclusions as well as a tongue thrust. We will use 100 children in each group. The list of boys will be assigned randomly to either the group with which we do therapy (the experimental group) or to the group which receives no therapy (the control group). We will make similar group assignments for the girls.

Now if you have been reading closely the information I have given you thus far, you may ask me, "how do you know the children are tongue thrusters?" The answer to this is determined by what criteria the researcher wishes to use and how he wishes to classify his subjects (in this case the children in the study). Therefore, he creates an operational definition of tongue thrust as well as class I open bite and selects his subjects (the children in the sample) on that basis. Subject selection takes many forms too. Most researchers will establish the criteria they wish to use first and then train observers (examiners if you will) in the use of the criteria until they reach a high level of agreement. This tells the researcher that whenever any examiner looks at a subject he will do so using the same criteria as any other examiner. This helps to reduce the bias which could occur if just one examiner looked at all the children and, of course, speeds up the selection process.

You may also ask another question. How will open bite be measured? The answer—anyway you wish it to be measured. For our study lets just take a ruler and measure the distance between the incisive edge of one lower central incisor to its maxillary counterpart. This will keep our study very clean (and keep me from being any more confused than I am).

We are almost ready to begin our experiment. There are just a few points we need to consider. One is the sample size. Is 100 children enough? Ordinarily, yes. But this depends upon how you wish to use the data. If we were using this study to determine the effectiveness of the polio vaccine where life and death or crippling illness was the issue, 100 children would not be enough to generalize for the entire nation. So you see, the size of the sample does enter into a decision of experimental design when the researcher must have data which may be generalized to the population he is studying. By-and-large, behavioral research does not deal with life-death issues. Behavior is more variable than most biological research. Therefore, behavioral scientists must accept lower significance levels and often smaller samples. Does this mean that behavioral research is inferior? No, only different.

Another point which must be considered is the therapy our experimental subjects will receive. Here we must decide on what kind they will get, how often they will get it, how long it will take per session, how long the entire program will take, who will do the therapy, and when we will measure the dental occlusion again to see if any change has occurred. Now I know you are waiting for me to tell you which therapy is best and how long it lasts and so forth. But for our purposes here we are not concerned with these value judgments. Value judgments have little place in research. What we need here are operational definitions and procedures which are consistent among the clinicians who will be doing the therapy. But let us generalize just a bit and say that the experimental subjects will be seen for three months of therapy followed by six months of follow-up checking. Following this nine month period we will wait another nine months to re-measure the dental occlusion as we did when we first examined our subjects. Thus, the entire experiment will take about 18 months. We won't discuss the specifics of our therapy program except to say that all experimental subjects receive essentially the same therapy and all subjects in both the experimental as well as the control group receive no orthodontic help or other forms of help which may contaminate our data.

Let's assume for a moment that we have completed the therapy, the six-month check-up period, the nine-month waiting period, and the measurement of the dental occlusion. We are now ready to analyze the results.

Before we get into a discussion of results, I believe we should briefly discuss two terms. One is "independent variable" and the other is "dependent variable." These are confusing terms to some people. But I believe you will find them easy to understand if you remember that the independent variable in a study is the one used to manipulate or change the behavior of the subjects and the dependent variable is the actual change that occurs. In our study, for example, the independent variable is the therapy the children get. The dependent variable is the change (or lack of it) which occurs in the dental occlusion. Now to the results.

In our little study we are interested to see if the therapy influenced the size of the open bite. We matched the children only insofar as sex is concerned, the presence of the open bite (not its size), age, and tongue thrust. Therefore, when we analyze our results we may make comparisons between all of the experimental and control subjects, between the girls, between the boys, or compare the boys to girls. In other words we may look at the differences in the open bite which has occured between the experimental and the control group of boys. We may also do the same thing for the girls. We may lump them all together to see if differences exist for boys and girls in the experimental and the control group. Finally, we may look to see if differences exist between the boys *and* girls in either the experimental or the control group. Suddenly our simple experiment becomes more complex, doesn't it? Let's see if we can clarify it for you just a bit. In our experiment we have controlled the variable of subject sex. We want to compare girls to girls and boys to boys. We took pre and post therapy measurement of the subjects dental occlusion and by simple subtraction determined the changes which have occurred for each subject. If an open bite has become larger, we express this as a negative number. Therefore, we have an arrangement which may look something like this (see figure 3). Now if we subtract therapy mean dental occlusions from the pre therapy mean dental occusions for each group, we may then compare the differences in the change between the groups of girls. We may do the same thing for the boys. We may also lump the boys and girls together and do the same thing. Now here is an advantage which should become apparent to you. It is possible that although a small difference was found in both groups (boys and girls) it was not sufficient to be significant. But if the sample size is doubled as in the case of lumping boys and girls together, those differences may then become significant. What we have lost is the ability to identify changes due to the sex of the subjects. Remember that the larger the sample is the more reliable the results become—all other things being equal. It will look something like this (see figure 4). Now if the data look as though there will be differences between the boys and the girls we may wish to see if these differences are significant. Therefore, we may wish to compare the mean changes in dental occlusion occurring among the boys with those occurring among girls: This is not a part of our original intent, but if the data are there, we should examine them for our data we might use a statistical test called the "t" test or we might use a simple analysis of variance or "F" test.

Girls - N = 50

	Experimental	Control
Pre Therapy	means of dental occlusion opening	means of dental occlusion opening
Post Therapy	means of dental occlusion opening	means of dental occlusion opening

Fig. 3. A way in which the results of the study may be displayed reading for analysis. In this case the results we obtained for the girl subjects.

Int. J. Oral Myol.
July, 1976

All - N = 100

	Experimental	Control
Pre Therapy	means of dental occlusion opening	means of dental occlusion opening
Post Therapy	means of dental occlusion opening	means of dental occlusion opening

Fig. 4. All subjects of the study may be used for analysis. Compare to Figure 3. The total "N" is doubled.

The way in which the data of our study are analyzed may vary according to what we want to know. Statistical analyses vary depending upon the size of a sample, the type of measurements we have used, and the distribution of the population we are studying. Statistics is a complex field of its own and is beyond the purpose of our discussion here. But we should say something about statistics generally so that you may understand some basic notions about it.

We have discussed some statistical terms in this paper already. Such terms as probability, level of significance, means, medians and modes are statistical terms. There are a few others which may help you understand a research study a little better.

Statistical measurements fall into two broad categories. One of these categories is where the data being analyzed may be expected to be distributed normally to conform to a bell shaped curve if the entire population was studied. The statistics used to analyze this data are called "parametric statistics." That is, we are studying a population for which we know the boundaries (parameters) and which is distributed in a way where most of its members will be found in the middle of the distribution. Human intelligence is such a phenomenon. That is, if we were to look at enough persons we would see that there are a few retarded persons and a few bright persons, but most of us would not be particularly retarded or bright. Lets look at a normal curve again to see how a normal curve is distributed (see figure 5). Notice that we have introduced two new terms in this figure—range and standard deviation when we look at a normal distribution curve as we have here. The difference between the lowest and the highest scores is the range. This is one measure of variability. Another is the standard deviation which is a more precise measure. A standard deviation is a method of describing differences of variability from the means of all scores falling within the range. Therefore, we may look at any score within the range to determine how far it differs from the mean of all the scores. You may ask, "so what?" The answer to this question is that by knowing where a score is in comparison with the other scores in a distribution, we can tell if it is a score we might expect to occur

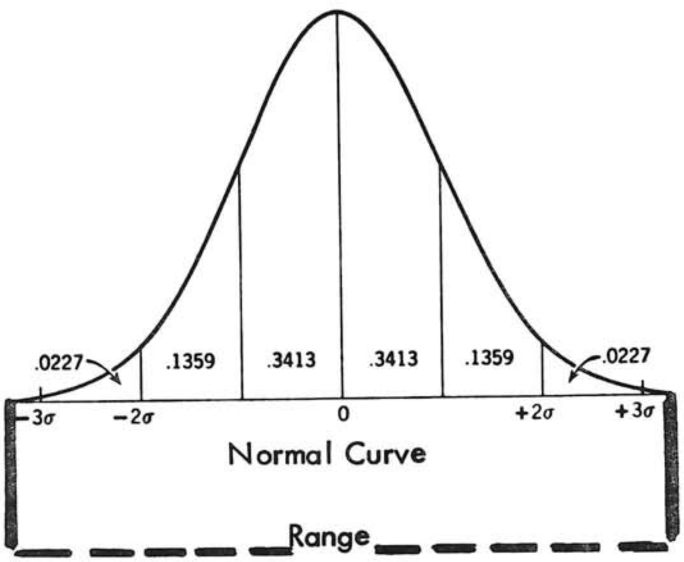

Fig. 5. A normal curve showing the range of the distribution and the percentages of the population which will be found within standard deviation.

frequently or one which is apt to occur only rarely. In intelligence testing, for example, when we have a score of 140 where the range is 200, the mean is 100, and the standard deviation is 20, we know that the subject scored two standard deviations above the mean. Now if you look at the bell shaped curve and you assume that intelligence follows a normal distribution curve, you realize how high that score of 140 really is. In other words, the kid with the 140 I.Q. is a pretty smart cookie!

The types of parametric statistical measures include: means, standard deviations, Pearson product—moment correlations, "f", "t", and "z" statistics and similar statistical computations which involve a population distribution.

The second type of statistical category is called "non parametric" statistics. Sometimes these are called "distribution free" statistics because they make no assumptions about the data falling into a normal distribution curve. Non parametric statistical measures include: modes, frequencies, contingency coefficients, medians percentiles, Spearmen correlations, Kendall correlations, Kendall coefficients of Concordance (w) and the X^2 test. Non parametric formulae may be used with data which conform to parametric criteria, that is, fall into a distribution curve. But since it is a general rule that parametric tests, this practice is used very seldom.

Summary

To summarize briefly, then, there are three basic forms of research designs: historical studies, descriptive studies and experimental studies. When these methods are compared, there is no "better way" to do the research, only different ways. Research is not a mysterious process to be feared by the reader or used as a means of intimidating him. It is used merely as a way to organize data and discover new informational horizons.

I would like to leave with you two thoughts which should prove helpful in developing a healthy attitude toward research. First, trust your own empirical observations. When you

Int. J. Oral Myol.
July, 1976

see something in your office time and time again which disagrees with the research you are reading, question that research. Probe deeply into its methodology and its design to see if you may identify weaknesses in it. Second, I would like to tell you an incident which happened to me which has given me "food for thought" over the years. As a doctoral student I was reading some research in my minor area, Educational Psychology, preparing a term paper for a most brilliant and world reknown professor. I attacked the research I had read with vigor and was incisive in my ability to discover the weaknesses of the research I was reading about. I turned in my paper which was read by this astute and gentle professor. He may have had several comments to make about the paper I had written, but I recall just one—the most important single comment I have ever had about any paper I have written. It said simply, "You have not given enough credit to the research process." Basically what this means to me is that any research you review will have its weaknesses. Any problem worth studying may be viewed from a number of different viewpoints. But this does not negate the research. It merely points to additional research which is needed. This is the very nature of research. As with your empirical observation, when several studies arrive at the same general conclusions, look hard at the data. If your observations differ from the research findings, question your observations. After all, this is the clinical method, isn't it?

REFERENCES

Auer, J. J., *An Introduction to Research in Speech*. New York: Harper and Brothers, Publishers, 1959.

Edwards, A. L., *Statistical Methods* (2nd Edition). New York: Holt, Rinehart and Winston, Inc., 1967.

Garrett, H. E., *Elementary Statistics*. New York: Longmans, Green and Company, 1956.

Hillway, T., *Introduction to Research*. Boston: Houghton Mifflin Company, 1956.

NcNewar, Q., *Psychological Statistics*. New York: John Wilen and Sons, Inc., 1963.

Runyon, R. P. and A. Haber, *Fundamentals of Behavioral Statistics*. Reading: Addison-Wesley Publishing Co., 1967.

Siegel, S., *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill Book, Inc., 1956.