

## **Understanding Central Tendency**

***Wannapa Trakulkasemsuk***

*King Mongkut's University of Technology Thonburi*

### **Abstract**

A measure of central tendency seems to be a frequently observed statistical method in research in Applied Linguistics. Theoretically, three most common measures of central tendency include mean, median, and mode. All the three measures are valid and said to be appropriate for different types of scales and/or conditions. Since research in Applied Linguistics covers a large area of studies, regular use of all these three measures should be assumed. However in most of research reports in Applied Linguistics which include a description of central location of data, the mean seems to be an only method found to be used. One may question if this occurs because of the nature of research in the field or the researchers' knowledge and familiarity of the measure. Thus, this paper aims to investigate how researchers in Applied Linguistics understand and select their measure of central tendency. Findings and discussion from this study may remind researchers about their selection and application of this very simple statistical method to strengthen their data analysis.

## **1. Introduction**

### ***1.1 Qualitative VS Quantitative research paradigm***

In general, research methodology can be classified into two main approaches namely qualitative and quantitative research. The two research approaches can be distinguished by their purposes, methods of data collection, types of data, methods of data analysis, and presentations of findings. Qualitative research deals with qualitative data such as data from interviews, narratives, phenomenologies, ethnographics, grounded theory, and case studies (Creswell, 2003; Dörnyei, 2007). The data of the qualitative research are non-numerical and thus statistical analysis is not required. Bordens & Abbott (2008) mention that although qualitative research can sometimes incorporate some quantitative data, the quantities included are no higher than ordinal scales (see section 1.3 for explanations of measurement scales). Thus, it does not require much or serious statistical operation. As opposed to qualitative research, quantitative research deals with quantitative data such as data from test scores, survey questionnaires and experiments (Creswell, 2003; Dörnyei, 2007). The data of quantitative research are numerical. Therefore, statistical methods are normally employed in the process of data analysis and interpretation.

The two research paradigms have their own strengths and weaknesses. The main strength of qualitative research is that it provides in-depth explanations of the data. However, since it analyses things in-depth, it usually includes a small sample size. As a result, it lacks generalisability. Moreover, qualitative analysis relies much on the researcher's judgments. Thus, it can be questioned about subjectivity. Quantitative research, on the other hand, requires a large amount of samples so findings from quantitative analysis can be generalised and said to be more objective since the results are derived from systematic statistical manipulation. Nevertheless, findings obtained from statistical analysis are mostly averaged. Therefore, it is possible that they may leave out some minor points which can be important or interesting.

To resolve the weaknesses of the two research approaches, mixed-method research, a combination of qualitative and quantitative approach, is suggested. There are several ways to conduct mixed-method research. For example, it can start with qualitative analysis to obtain some qualitative findings. Then, the qualitative findings are quantified in order to generalise

the results. It can also be done in the other way round that is to begin with quantitative analysis to get a list of quantitative findings. Then, detailed qualitative exploration can be done to describe the quantitative findings. Furthermore, the two types of data (i.e. qualitative and quantitative data) can be collected and analysed concurrently so as to triangulate the results (Creswell, 2003).

Research designs in Applied Linguistics can be based on any research approaches. However, a study of current trends in research in Applied Linguistics, by reviewing papers published from 1991 to 1997 in good quality international journals, has manifested a significant higher proportion of quantitative research studies over qualitative ones (Lazaraton, 2000). Until recently, the same trend can be widely observed. Most of research studies in Applied Linguistics, to a certain extent, usually incorporate, at least some, quantitative methods.

Since quantitative research seems to be a major trend in Applied Linguistics, researchers in the field, even though those who favour qualitative over quantitative approach, may have to be aware of it when designing their research or reading other people's research reports. Certainly, some statistical methods, either simple or complex, must be employed in quantitative studies. Sound understanding of basic concepts in statistics can be beneficial.

### ***1.2 Descriptive statistics: the basic and simple statistical method***

There are two principal kinds of statistics which are descriptive and inferential. Descriptive statistics is used to describe or summarise general behaviors of the whole set of data. After conducting a descriptive analysis, if researchers aim to make further inferences or predictions of the entire population from the samples, inferential statistics can be done.

Descriptive statistical methods are all very simple to understand and to compute. It only requires basic mathematics knowledge and operations (i.e. counting, adding, subtracting, multiplying, dividing etc.). As it is easy and is always required in any quantitative analysis (while inferential statistics can be later or optional), it is possible and (can be) necessary for researchers or readers of research reports to understand it well.

Descriptive statistics includes several topics such as graphics, frequencies, percentages, central tendency, skewness, standard deviation and so on. Among these, central tendency tends to be one of the most frequently topic observable in Applied Linguistics research reports. Although the calculation and interpretation of central tendency is very easy, I do have a frustrating experience with the use of it. That is since there are more than one way to obtain the central value of the data, applying a less appropriate procedure may lead to a distortion of the true behavior of the whole data set. Therefore, the aims of this paper are first to discuss the concepts and share experience in using central tendency and second to investigate how researchers in Applied Linguistics understand the concepts and how they use it.

### ***1.3 Measurement of scales***

Before getting to further discussion of the central tendency, scales of measurement is a crucial concept that needs to be discussed because they are factors determining what statistical test should be used. Scales of measurement are used to distinguish variables. There are four types of statistical scales namely nominal scales, ordinal scales, interval scales, and ratio scales.

*Nominal scales* are used to label variables. Numerals are assigned to variable categories only for the sake of categorisation. They do not carry any quantitative value. For example, to label occupations of research participants, a researcher may assign '1' for 'teacher', '2' for 'engineer', '3' for 'doctor' and so on. In this example, no quantitative value

can be interpreted. Therefore, a statement like ‘doctor’ (3) is greater or better than ‘teacher’ (1) cannot be made.

*Ordinal scales* are used to represent the rank ordering of variables. Thus, the numbers of this type of scales are said to have quantitative meanings. An example of this scale type is a set numbers derived from a rating scale. For instance, a researcher may ask participants to rate their agreement on a particular issue, so ‘1’ may be assigned for ‘disagree’, ‘2’ for ‘neutral’, and ‘3’ for ‘agree’. Thus, participants with a 3 do agree on the issue more than participants with a 2 or a 1 respectively. Although ordinal scales do have quantitative values in terms of ranking (e.g. 2 is greater or higher than 1), they cannot truly specify how much the values of different ranks differ. Therefore, an interval between ‘1’ and ‘2’ and an interval between ‘2’ and ‘3’ may not be equal. Hence, decimal numbers like ‘point five’ (.5) in 1.5 and 2.5 may not always equal. In fact, decimal scales of ordinal variables should not be computed.

*Interval scales* are quite similar to ordinal scales in that the numbers are quantitative. The difference between ordinal and interval scales is that interval scales represent an equal interval between each pair of consecutive numbers. Examples of interval scales are measurements of the length of time and temperature.

*Ratio scales* have all the properties of interval scales. The difference between the two scales is that ratio scales have a value of an absolute zero. Ratio scales are said to be rather common in the field of physical sciences than in psychology (Ferguson & Takane, 1989). Examples of ratio scales are measurements of weight, height, and distance.

It should be noted that the distinction between interval and ratio scales are not always clear-cut since it is sometimes difficult to state if there is an absolute zero of the scales. Take the test score as an example. It is classified as an interval variable in some statistical textbooks (e.g. Hatch & Lazaraton, 1991), but as a ratio variable in some other textbooks (e.g. Wrench et al., 2008). Fortunately, most of the statistical methods treat the two types of scales in the same way. Thus, the two types of scales can be analysed with the same statistical tests. In most computer programs for statistical analysis like SPSS, for example, interval and ratio scales are included under the same variable type.

#### **1.4 Central tendency<sup>1</sup>: Mean, Median, and Mode**

Measure of central tendency or central location is a statistical concept for typifying the whole set of data. In other words, it explains the global behavior of all samples in an average manner. The most common types of central tendency are *Mean*, *Median*, and *Mode* (Brown, 2004; Ferguson & Takane 1989; Hatch & Lazaraton, 1991).

##### **1.4.1 Mean**

The arithmetic mean or (generally referred to as) mean is claimed to be the most widely used method to describe central tendency (Dörnyei, 2007; Ferguson & Takane, 1989; Porte, 2002). It is described to be a method of central tendency which can offer a more accurate or more efficient estimate of the population than other measures i.e. median and mode, as it takes all values in the data set into account (Ferguson & Takane, 1989). The mean is calculated by adding all the scores in the data set, then dividing by the total number of the scores.

For example, we have test scores of ten students: four students got 15, two 16, two 17, one 19 and one 20. The mean of this group of students can be calculated as follows.

$$\text{mean} = \frac{15+15+15+15+16+16+17+17+19+20}{10} = 16.5$$

Although the mean is said to be a powerful method to predict the central location of the data set, it should be noted that the mean is most appropriate only for interval or ratio scales (Ferguson & Takane, 1989; Spatz & Kardas, 2008). In other words, when we have nominal or ordinal scale variables, the mean may not explain the average of the population properly. For example, when we assign '1' and '2' to label gender categories of 'male' and 'female' respectively. The calculated mean of 1.3 cannot describe anything about the average gender of the participants.

Another point that researchers should be aware of when using the mean is that the mean can best portray the population when the distribution is normal. This is because the mean represents the centroid (center of gravity) of the data so it is rather sensitive to the outliers especially when the sample size is small. Consider this case as an example. A list of 5 students' test scores includes 99, 54, 50, 42, and 38. The calculated mean of this data set is 56.6. We can see that the mean is higher than the real scores of four students. Thus, this is clear that the mean cannot truly explain the phenomenon since it is skewed by the outlier, 99. To solve this limitation of the mean, another measure of central tendency like 'median' should be considered.

#### *1.4.2 Median*

The median is another commonly used measure of central tendency. It is the exact middle point of the data. Half the scores in the data set are below the median, and half are above. To obtain the median, first all scores are arranged in rank order, usually from low to high. Then, the middle score can be identified. If the total number of the scores (N) is an odd number, the median can be directly stated (e.g. N = 5, the median = the 3<sup>rd</sup> rank score). If the total number of the score is an even number, the mean of the two middle scores is to be computed (e.g. N = 6, the median = (the 3<sup>rd</sup> rank score + the 4<sup>th</sup> rank score)/2).

The median is appropriate for ordinal scales (Ferguson & Takane, 1989). Also, it can be used instead of the mean, when the mean is not appropriate or the sample size is small (Hatch & Lazaraton, 1991; Porte, 2002). Turning back to the above example when a set of data includes five scores: 99, 54, 50, 42 and 38. The median obtained from this data set is 50, while the mean is 56.6. In this case, we can clearly see that the median can offer a better summary of the data than does the mean.

However, the main weak point of the median is that it only concentrates on the middle score(s) and thus disregards others. Therefore, it can sometimes be doubtful about its power to really represent the whole data set. Consider these two data sets as an example; Set 1: 99, 54, 50, 42 and 38 and Set 2: 99, 60, 50, 48 and 46. The modes of the two sets are the same (50) although the distributions of the two sets are totally different.

A technique called 'trimmed mean' is suggested as a compromise between the mean and the median, when both of them cannot very well describe the central tendency (Larson-Hall & Herrington, 2009). The trimmed mean is a technique to deal with the outliers or extreme scores and the exclusion of most scores in the data set. To do so, it trims off the scores on the two ends of the data set then calculates the mean. To trim the scores, Wilcox (2010) suggests that 20% is a good amount for the trim of each end. The same sets of data above can be used again to illustrate the technique. When N = 5, 20 per cent of it means 1. As a result, we have three remaining scores in each of the data sets which are 54, 50, 42 for set 1 and 60, 50, 48 for set 2. The trimmed mean of set 1 is 48.7 and set 2 is 52.7. These two figures can then offer a better explanation of the data sets.

One more point to note is that, even though it is highly suggested that the median is the most appropriate measure for ordinal scales (i.e. scores obtained from rating scales), the mean can also be applicable if an equal interval between scores in a rating scale and a normal distribution can be assumed.

### 1.4.3 Mode

The mode is a very simple method to indicate the central location of the data. It is the value that occurs most frequently in the data set. Because it offers general and less precise description, it is applicable to all types of scales including nominal scales which do not permit the mean and median. Furthermore, when the distribution is abnormal and the mean and the median cannot provide a good estimate, the mode can be used. Consider the following data set obtained from a five-point rating scale.

rating	1 (very disagree)	2(disagree)	3 (neutral)	4 (agree)	5 (very agree)
frequency	30	14	14	13	30

This data set is bi-modal; it has two modes which are 1 and 5. If a researcher of this study blindly interprets the data based on the median (seeing the data as ordinal scales) or the mean (assuming equal intervals), the values of the central tendency of this data set are 3 and 3.02 respectively. The median and mean scores here lead to misunderstanding of the averaged responses from the participants by concluding that the participants have neutral attitude in general. In this case, it is obvious that the mode is the best method to describe the phenomenon.

Nonetheless, there is a case where the mode cannot be found for example when all nominal categories obtain the same frequencies. In this case, a more simple descriptive statistics like *percentage* is more appropriate (e.g. 50% of the participants are male and 50% are female).

To conclude this section, I would suggest that central tendency is a nice statistical method to typify and describe general characteristics of research findings. Each of the measures of central tendency has its own strengths and weaknesses. Researchers may need to look at their data carefully before selecting what method to apply. From my observation, most researchers I know, including me myself before facing a severe case in one of my research studies, usually take it for granted and use the mean every time they want to summarise the data. To deal with central tendency more wisely, I would rather suggest that researchers should run all the three measures (if all applicable based on the types of scales) and compare the values. Since we have computer programs for statistics, this can be done very easily. Furthermore, I would as well suggest researchers to consider graphic representation of the data as it can help support a better explanation of statistical values.

As I am curious to know how teacher-researchers<sup>2</sup> in Applied Linguistics understand and use the measures of central tendency in their research studies, I also investigate it. The sections below present research method, findings and discussion.

## 2. Method

### 2.1 Subjects

The subjects of this study were 7 Thai teachers of English at King Mongkut's University of Technology Thonburi (KMUTT). Even though the seven subjects' main job was a teacher, they were selected as subjects of this study because all of them had sound experience in doing research in Applied Linguistics. Each of the subjects was invited for a semi-structured interview. The duration of each interview ranged from 15 to 30 minutes. The interviews were all done in Thai.

### 2.2 Instrument

An instrument used in this study was semi-structured interviews. Four main guided questions and eight follow-up questions were prepared. The participants were asked about

their attitudes towards the use of statistics in research, their experience of using statistical methods in research studies, and their knowledge of central tendency. After the elicitation of the subjects' understanding of central tendency, the researcher explained the concepts and procedures of the mean, mode and median to the subjects. Then, the subjects were asked to discuss their thoughts about the applications of the different measures. The interviews were conducted in a very informal manner. During the interview, the subjects were allowed to ask questions, argue, or express their thoughts freely.

Information from the interview of each subject was recorded both in the written forms and audio files. To conclude the findings, the two forms of data were cross checked so as to prevent any missing of the information.

### **3. Findings and discussion**

#### ***3.1 General attitudes towards the use of statistics in research studies***

Five out of seven teacher-researchers mentioned that they felt totally uncomfortable when they had to use statistics in their research studies. They were uncertain about their knowledge and ability to select appropriate methods. If possible, all of them admitted that they tried to avoid using it. However, all of these five subjects, though they felt rather negative, accepted that an integration of statistical methods to research data analysis was useful.

The other two subjects said that it was fine with them to include statistical analysis in their research studies. Certainly, they felt more comfortable with the methods that they were familiar with. They also expressed their awareness of the selection of a proper statistical method. Both subjects explained that when they needed to use some new methods, they had to study them carefully.

When asking about how statistics useful for research, all the subjects replied in the same positive direction. All believed that statistics could make it easy for them to conclude, interpret, and generalise the findings. They stated that statistics helped provide solid and reliable presentations of findings. Also, with statistical results or scores, they thought they could explain their findings more easily.

The findings on the subjects' attitudes towards the use of statistics in doing research reveal some conflicts between the subjects' views on the difficulty of applying statistics in research and its usefulness. They saw it difficult when they had to select and run the tests. On the other hand, after they got the results, they found it easy for them to report the findings.

According to this, I would say that the teacher-researchers interviewed in this study did not really have negative attitudes towards the use of statistics in Applied Linguistics research. In fact, they perceived it as a useful thing to include in their studies. The real problem might be that most of them lacked concrete knowledge and confidence.

To encourage teacher-researchers in Applied Linguistics to use more statistics, ones may need to think of how to enhance their knowledge of statistics and confidence. Consultations and knowledge sharing channels should be worth-considering. At this stage, I do not want to make a suggestion that the researchers should study more solely by themselves. This is because most of statistics books are not written in a reader-friendly manner. Thus, they might demotivate or discourage the researchers who do not have good enough background knowledge. Moreover, I do not think that statistics training can really help. Researchers with different research questions and designs need to know different statistical methods. A general training session to everyone may not be as helpful as an individual consultation. However, it seems to be impossible to get a consultant who can answer everything. Thus, I do believe that having a channel or a community where researchers can discuss their problems, share their knowledge and experience, and/or help each other to study and find solutions may be the most effective way.

### **3.2 Knowledge and use of 'central tendency'**

Unexpectedly, all the subjects did not know the term 'central tendency' at all. When they were asked to guess or give some examples, all of them only literally translated the term into Thai.

Then, I told the subjects that mean, median, and mode are the three common measures of central tendency. All of them said that they knew what the mean was and they always used it when they wanted to present an average of the research samples. Some of them mentioned that it is used in grading students' performance at the end of the semester. When the subjects who mentioned the use of the mean in grading were asked further about how they interpreted the mean scores when grading students, none of them managed to explain. Apart from the mean, only four of them knew what the mode was, but have never used it. Six of them said that they did not know the median at all and thus have never used. The only subject who knew the median reported that she just knew what it was, but did not know how to use so she had never used it.

After that, an explanation of the three measures was given to the subjects. They were allowed to ask for clarification. Then, I asked them to discuss how, when, and which kind of data the three measures should be most appropriate for. All the subjects gave the same answer that the mean can be used with nearly all types of data. The examples they gave were also the same. They mentioned data from test scores and rating scales. As for the median, none of the subjects had an idea about its application. Like the median, all the subjects were uncertain about the use of the mode. Only two subjects tried to guess that it should be used to explain the frequency distribution.

At the end of the interview, each subject was asked to conclude the concepts of the three measures or provide his/her reflections. Four subjects did not discuss this point. The other three subjects who responded to this prompt provided different answers.

The first subject stated that the three measures should serve different purposes and should be used under different conditions. However, she said she was not very certain about their different purposes and conditional requirements. She also mentioned that she always used the mean because it was most familiar to her, and never used other measures to report her research findings.

The second subject expressed that she always used the mean to report the central tendency in her research studies. The median and mode were new to her. However, she felt fascinated to know them. She said that she would be more careful about the selection of the measures in her future research.

The last subject did not provide any conclusion of the concepts. He only reflected that he did not care much about it because he was more on the qualitative side. The only statistical method he usually employed in research studies was the percentage. In addition, he said that if he happened to do any research studies that involve statistical analysis other than the percentage, he would pair up with someone who know it and let the co-researcher work out the statistical analysis while he would be responsible for something else.

Findings of this section reveal that the subjects did not clearly understand the concepts and the measures of central tendency. It seems that the mean is an only method they recognised and used because it is familiar to them. For all the subjects, the mean seems to be an only method available for the measurement of central location. This understanding could lead to a danger of overgeneralisation of the use of the mean since the subjects did not explicitly show their awareness of the limitations or misuses of the mean.

That no one mentioned anything related to the scales of measurement at all can be interpreted that the subjects did not know or were not aware of it. As the scales are important in determining statistical methods, a lack of knowledge of the scales may result in an

inappropriate selection and use of the methods. Consequently, research studies with inappropriate statistical analysis may provide misleading interpretation and conclusion.

It is common for the nature of research in Applied Linguistics to display an average figure or central tendency of the data. That the mean is regularly found in research reports in the field is also common since the mean is claimed to be the most powerful measure comparing to the other two. However, researchers should use it because it is really appropriate, not because it is the only method they are familiar with. To select the measure of central tendency, first the researchers should know what type of scales their data is. If it is nominal, the only choice is the mode. If it is ordinal, the mode and the median can be used. Last, if the scale is interval or ratio, all the three measures are applicable. When two or more measures are possible, researchers may have to test all of them and compare the results in order to choose the best one to explain the data. It should be noted that when findings are presented in a research paper, readers cannot see the raw data and they only learn from what the researchers present. They hardly know if the finding is misleading (e.g. using the mean for bi-modal data as the above example). Even though this kind of mistake is hard to detect without seeing raw data, it is the researchers' responsibility to choose the best method to obtain the most reliable presentation.

#### **4. Conclusion**

Statistics is said to be a tough subject to most of Thai researchers in the field of Applied Linguistics. I admit that I may over claim this. However, based on my direct experience as a student, researcher, teacher and whatsoever in the field, I always hear my colleagues' and students' complaints about it. In fact, statistics does not have to be very complex. Like central tendency, it comprises very general notions since people, by their nature, usually average things around them. In addition, the calculation procedures of all types of its measures are not at all difficult. This paper may, hopefully, facilitate more understanding of central tendency to people in the field. Also, it may alarm researchers to be more aware of their use of statistics, even the simple and very familiar ones like the mean.

This study includes only a small number of subjects from only one context. Thus, the findings may not be generalisable. However, since the subjects were selected randomly from a group of teacher-researchers, findings from this may be able to represent some nature of researchers in Applied Linguistics.

#### **Notes**

1. Any measures of central tendency should be reported with the dispersion of the frequency distribution (e.g. range, variance, standard deviation). However, explanations of such things are beyond the scope of this study. Researchers who aim to include central tendency in their research report are recommended to further study it.

2. I use the term 'teacher-researcher' to notify that the subjects included in this study were not a hard-core type of researchers. In fact, they were university teachers who regularly do research.

#### **References**

- Bordens, K. S. & Abbott, B. B. (2008) *Research Design and Methods: A Process Approach*. New York: McGraw-Hill.
- Brown, J. D. (2004) Resource on Quantitative/Statistical Research for Applied Linguistics. *Second Language Research*, 20(4), 372-393.
- Creswell, J. W. (2003) *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. California: Sage Publications.



- Dörnyei, Z. (2007) *Research methods in applied linguistics: Quantitative, qualitative and mixed methodologies*. Oxford: Oxford University Press.
- Ferguson, G. A. & Takane, Y. (1989) *Statistical Analysis in Psychology and Education*. New York: McGraw-Hill.
- Hatch, E., & Lazaraton, A. (1991) *The research manual: Design and statistics for applied linguistics*. New York: Newbury House.
- Larson-Hall, J. & Herrington, R. (2009) Improving Data Analysis in Second Language Acquisition by Utilizing Modern Developments in Applied Statistics. *Applied Linguistics*, 31(3), 368-390.
- Lazaraton, A. (2000) Current Trends in Research Methodology and Statistics in Applied Linguistics. *TESOL Quarterly*, 34(1), 175-181.
- Porte, G. K. (2002) *Appraising Research in Second Language Learning: A practical approach to critical analysis of quantitative research*. Philadelphia: John Benjamins.
- Spatz, C. & Kardas, E. (2008) *Research Methods: Ideas, Techniques, & Reports*. New York: McGraw-Hill.
- Wilcox, R. R. (2010) *Fundamentals of Modern Statistical Methods: Substantially Improving Power and Accuracy*. New York: Springer.
- Wrench, J. S., Thomas-Maddox, C., Richmond V. P. & McCroskey J. C. (2008) *Quantitative research methods for communication: A hands-on approach*. Oxford: Oxford University Press.