PART 2
An Experiment on the Box plot

Mahmud M. Hussin
Statistics Dept. College of Admin. & Econ.
Baghdad University

الخلاصة

البحث يحتوي على تجربتين لاجراء بعض التعديلات على الرسم الصندوقي (box plot) . هذه التجارب تم تصميها لأيجاد تأثير بعض التحويرات (تغيير العرض للصندوق (box plot width) وكذلك تغير الامتداد الخارجي للصندوق (whisker length) المحددة في الرسم الصندوقي .هذا الشكل أو الرسم أو الادات تعتبر من أفضل وأهم طرق الرسم على الاطلاق لعرض البيانات وكذلك لاجراء المقارنة بين عدة مجاميع وحسب الغرض المطلوب من البيانات ممكن المقارنة لجميع خواص العينة أو لوحدة من الصفات والرسم يحدد خمس قيم أساسية من البيانات وهي أكبر قيمة وأصغر قيمة والربيعات الثلاث. الاشخاص الذين ساهموا بهذه التجربتين طلب منهم أن يعملوا مقارنات في كل مرة بين زوج من الرسم الصندوقي (box plot) يكون واحد ثابت لجميع المقارنات(standard plot) ويكون في ورقة منفردة والثاني هو المتغير ويكون كل تجربة في ملف تمثل ٤٥ شكل(تغيير الطول للرسم (box length) مع العرض (box width) هذه التجربة الاولى) و (تغير الطول للرسم (box length) مع الامتداد الخارجي(whisker length) في ملف ثاني (booklet)) كل شخص يساهم بتجربة واحدة ويقارن أطوال الاشكال (box plots) في أحد الملفات مع الرسم الثابت (standard plot) في كلا التجربتين طلب من الاشخاص أن يحدد في كل مقارنة من ال ٤٥ مقارنة أياً من أطول الصندوقين أقصرلتحديد الاتجاه هل هو الثابت أم الذي في الملف ويخمن نسبة مئوية طول الصغير من الكبير وبشكل سريع أول تصور يلاحظه وليس قياس دقيق. التجارب بنية كأمتداد للتجارب التي قام بها (Hussin, M. M. (1989, 2006) عندما أقترح في الدراسات المستقبلية النقاط (1, 2) أن تغير الطول الى الرسم الصندوقي الثابت (standard box plot length) سوف يؤثر على أدراك الشخص من طول الصندوق (box plot length) عند المقارنة بين هذه الازواج من الرسوم الصندوقية (box plots). النتائج لهذه التجارب تؤكد على أن هذه التحويرات أو التعديلات تؤثر على أدراك الأشخاص لأطوال هذه الرسوم الصندوقية (box plots).

نعتقد أن هذا التأثير في قرارات الأشخاص لهذه التحويرات على ا لـ (box plots) ممكن أن تكون كنتائج للتفاعلات بين المتغيرات للرسم الصندوقي والتي هي (box length, box width, whisker length) وهذه التفاعلات تكون ما يسمى بالوهم المرئي (visual illusion) وهذا يؤثر على قدرة الأشخاص لأتخاذ قرارت دقيقة بالنسبة الى أطول الرسم الصندوقي (box plot lengths) . وكذلك النتائج تؤكد على أن الاقتراح بتغيير الطول الى الرسم الصندوقي الثابت (standard box plot length) في الفقرات (١ و٢) للدراسات للمستقبل كانت صحيحة. هذه التجارب أجريت على طلبة قسم الاحصاء المرحلة الثالثة والرابعة/ كلية الادارة والاقتصاد /جامعة بغداد.

Abstract

These experiments seek to investigate the effects of the fixed

variations to the basic box plot on subjects' judgments of the box lengths. The study consists of two experiments, were constructed as an extension to the experiments carried out previously by Hussin, M.M. (1989, 2006). Subjects were asked to judge what percentage the shorter represented of the longer length in pairs of box lengths and give an estimate of percentage, one being a standard plot and the other being of a different box length and also varying with respect to other elements such as, box width or whisker length. When he (1989) suggested in the future research points (1, 2), the changing length of the standard box plot effects on the subjects' perception of the box length. However, both experiments were used the stander box length as the middle box length levels in the experiments. The results of these two experiments indicated that these variations effected the subjects' perception of box length. we thought that the effect in the subjects perception of these variations it might be that the subjects were affected by the visual illusion effects as Cleveland et al (1987) accepted in their replies to the comments on their work, as a results of the interactions between box plot features as which effect the subjects ability to accurately judge box length and the effects differed between variations, both experiments were run in statistics department, Baghdad University.

١ -*Introduction.*

Graphical methods are an essential part of the exploratory data analysis. These techniques can give us a clear idea about the patterns of data set distributions, and can separates out elements of the data and reveals them to the data analyst. Also can help to show the unexpected features of the data or can allow us to make simple or detailed comparisons between distributions of the data sets. Graphical methods are used not only to summarize data, but also as diagnostic aids in analysis, and to decoding of quantitative information from the graphs. These tools represent a great part in exploratory data analysis in statistics, and have a long history of use in preparing pictures of data and presentation. Understanding the graphical methods will allow the data analyst to make graphs that transfer quantitative information to the viewers with more precision and effectiveness (Chambers et al, 1983).

The box plot (Turkey, 1977) is one of the important tools of the graphical methods. This tool can give the viewers a fast idea of fixed features of the distribution, the shape and the spread of the data. This tool can be applied to make simultaneous comparisons between the distributions of several sets of data. The idea of the box plot is simple (see example no.1), it is a graphical display which uses five values obtained from the data set, the upper and lower (hinges), the median, and the upper and Lower adjacent values. This paper is concerned to investigate whether fixed variations to

a box plot affect subjects' judgments of the box length (midspread). Features studied include making the box width proportional to sample size and the whisker lengths equal the adjacent values (Tukey et al .١٩٧٨) Subjects were asked to make comparisons between two box lengths and to give what percentage the shorter was of the longer. These experiments were carried out at Baghdad University. There were two statistical methods used to analyze the data of these two experiments, the analysis of variance techniques and the median polish techniques. The results of these experiments suggest that these two variations are affected subjects judgments of the box length.

٢ -*The earlier studies on box plot.*

Experimental work on graphical methods was conducted by Cleveland et al,(1982, 1984, 1986, 1987) and Simkin & Hastile, (1987), for detail of experiments and results see Hussin, M. M. (1989, 2006). Five studies have been carried out in box plot. In (1981), Mc Culloch run an experiment to study the effect of three variables, box length, box width, and viewing time on the subjects judgments of the box length. Mc Culloch concluded that the subjects reaction time of the box length affected by the two variables, box width, and viewing time. The interaction between two variables box length and box width affected subjects' judgments of box length. Box width added more information to the box plot but made the interpretation of the box plot more difficult, and for more detail see Hussin, M.M.(1989, 2006).

And in 1982, another study four experiments in box plot were run by Knight, to examine the effects of varying four features of a box plot on subjects' judgments of box length. These variables investigated in four separate experiments were box width, box notch, whisker length, and outlier values. Knight found that box width and box notch, affected the subjects' judgments of the box length, but the other two variables, whisker length and outlier values did not affect the subject judgments of box length. The outlier value, the observation their position is beyond the whisker length,and for more detail see Hussin, M.M. (1989, 2006).

The third study by Hussin, M.M. in 1989, investigated the effects of vary three features of a box plot on subjects judgments of box length, in two different groups of experiments and carried out at Keele and Baghdad Universities. In the group A experiments (comparative experiments), three experiments to study the effects of three variations to the box plot Tukey (١٩٧٧) on the subjects judgments of the box length, these variation box width, whisker length, and box width with box notch. Subjects were asked to make comparisons between pair of box plots, one of the pair is the standard plot and the other from the booklet. Subjects

were asked to respond if the box length of the box plot of the booklet is shorter or longer from the standard plot, and to give a rating of how confident in their judgments by giving a score of 50% to 100%. The subjects were asked to give their answers as a first impression. Hussin, M.M. (1989) found for the box width experiment at Keele, that there is a significant interaction between the two variables, box width/ length. The two variables box width and box length affected subjects judgments of box length, but the box length more than box width. And for the box width experiment at Baghdad found that the results are similar to the keele experiment results.

Hussin, M.M. found for the whisker length experiment at Keele, that there is a significant interaction between two variables box length and whisker length. These two variables are important and affected subjects judgments of box length, but box length more affected than whisker length. And for the Baghdad experiment found that there is no significant interaction between two variables, whisker length and box length. which is different from Keele experiment, Also these two variables are important.

For the third experiment of three variables box notch experiment he found that for Keele experiment, there is a significant interaction between these three variables of this experiment box width, box notch and box length, and only one interaction was significant of the two way interaction, box length and box width interaction. These three variables are important and affected subjects' judgments of box length. The subjects have more difficulty with judgments in this experiment than in all the other two factor experiments, make more error judgments, and have little confidence in their answers. He found for the same experiment at Baghdad that similar results for three way interaction was significant, but different for two-way interaction the box notch and box length was significant. And these three variables are important for the experiment.

In group B experiments (ratio experiments), which contains four experiments, two of them for length judgments and the other two for area judgments. Two of them were carried out at Keele and all of them at Baghdad. These experiments seek to examine, which features of box plot affected subjects judgments of box length (midspread). Subjects were asked to give percentages for how much shorter , smaller, represented of the longer or larger the length or area of box plot from the pair of box plots, one of the box plot being standard and the other with one or more of the features changed, these compared side by side on A4 sheet of paper. Their effects on judgments were estimated by the error size. Absolute value of the error = [judged percentage - true percentage].

The box length experiment at Keele, was build to study the effects of

the box length and whisker length variables on the subjects' Judgments. Hussin, M.M., found that there is interaction between these two variables. The box length variable is important and affected subjects judgments of box length, and more important than whisker length. The subjects tend to increase the midmeans of the absolute error with increase in box length, and similar results for the same experiment at Baghdad.

For the box plot three variables, box length, box width, whisker length, Baghdad length experiment. He found that, there are some interactions between these three variables, these interactions affected subjects' judgments of box length. The subjects in this experiment with change in three variables faced more difficulty than in any other experiments using two variables that would means these variables add more difficult to interpret the box plot, and affected the subjects' ability to make accurate judgments.

And for the box plot two factors area experiment at Keele, box length, whisker length, he found that, there is some interaction between these two variables, also the box length very important, but the whisker length less important for this experiment and this result for area experiment very reasonable the whisker length variable is not relevant, the subjects tend to overestimate with small areas and underestimate with large areas.

And for the same experiment at Baghdad. He found that, there is not significant interaction between these two variables, whisker length and box length. The box length variable is very important and dominated the experiment, by affected subjects of box length; the subjects tend to increase the errors with increases on the box area and also the same with keele experiment, this result for area experiment very reasonable and important the whisker length variable is not relevant.

And for the box area, three variables experiment, at Baghdad, Hussin, M.M. concluded that, there is significant interaction between these three variables, and also between any two of them, box length/width, and whisker length ; The two variables box length/width are important, and affected subjects judgments of box length, but whisker length was not important on its own in this experiment and also the same with previous area experiments this result for area experiment very reasonable and important the whisker length variable is not relevant.. Subjects faced difficulty with this experiment more than with other experiments and made large errors. Area judgments are more difficult than length judgments, these results agree with Cleveland & McGill (1984), and the power law results, and also weber's law might help to explain the results.

The fourth study by Sim, C.H.; Gan, F.F.; Chang, T.C. in 2005, they focus on the detection of possible outliers based on the box plot procedures. The outliers in a set of data are defined to be a subset of observations that appear to be inconsistent with the remaining observations. They indicate that the commonly constructed box plot is in general inappropriate for detecting outliers in the normal and especially the exponential samples. And they suggest that the graphical box plot be constructed based on the knowledge of the underlying distribution of the dataset and by controlling the risk of labeling regular observations as outliers.

The fifth study by Hussin, M.M. (2006), two experiments to investigated the effects of vary two features of a box plot on subjects' judgments of box length, these variation box width, whisker length and carried out at Baghdad University. Subjects were asked to make comparisons between pair of box plots; one of the pair is the standard plot and the other from the booklet. When in this study the standard plot (box length) is the smallest one of the box length levels in these two experiments, which is different from all these previous studies in the box plots, and also he found the results of these two experiments different from all these previous studies in the box plots the interactions between the two variables and box length very highly significant and the whisker length very important by it self and with the interaction and these two variables very important with the box length variable which this result different from the others and agree with the suggestions. When he (1989) suggested in the future research points (1, 2), the changing length of the standard box plot effects on the subjects' perception of the box length.

٣ *The problem suggested for this study*.

Now let try to explain the errors in subjects' judgments in interpreting the box plots found in previous studies of box plots, Hussin, M.M. (1989, 2006). It might be one possible explanation of the errors in the subjects' judgments of the box plot is that the interaction between the box plot features with the box length variable. The subjects might underestimate the box length when boxes are wider as changed the width or have longer whisker lengths and vice versa, this results might be similar to Baldwin's (1895) figures, when he found that the line lengths closer to the large square look shorter than that line lengths closer to the small square, for more detailed see Hussin, M.M. (1989, 2006).

Another possible explanation of the biases of subjects' judgments in interpreting box plots by using Cleveland & Mc Gill's (1984) theory. One possible interpretation might be that the subjects make area judgments instead of length judgments in the case of varying box width or box notch in combination with box length. And that make the judgments more difficult, because the area judgments more difficult than length judgments, power law (Stevens, 1975), and Celevland & Mc Gill (١٩٨٤).

Moreover, in the case of varying box notch with box length, the subjects might face difficulty because subjects need to make two length judgments, one for box length and the other for box notch, in addition to the interaction between these two variables. Also as Lovie (Lovie 1985) discussed the nature of the box plot is not a simple graph by which to make quick judgments for more detail see Hussin, M.M.(1989, 2006).

Another law also might help to explain the problem in judging box plot, in 1834, Weber proposed what we call now Weber's Law (Stevens, 1975) and we can give simple idea of this law is that when we need to make comparison between lengths of two things we need first to determine the difference between them by fixed percentage, and not on the overall sizes of the two lengths. Also Stevens (1975) proposed power law might help to explain the errors in the subjects' judgments of the box length of the box plot, and this law used to determine the accuracy in the judgments of different physical aspect objects, such as area, volume, Length, or... etc. The law state that the accuracy of these aspect judgments can be ordered as follow, length, area, and volume.

The problem suggested for this study is the investigation of the effects of certain variations in the box plot on subjects' judgments of box length. To find which of the two box plot features, box width, and whisker length effect on the subjects' judgments of the box length as a relevant factor. Subjects were asked to make comparisons between pair of the box lengths of box plots placed side by side; we tried to make accurate judgments and to avoid any effects of the orientation on the subjects' judgments. When in this study the standard plot (box length) is the middle one of the box length levels in these two experiments, which is different from all these previous studies in the box plots. when Hussin, M.M. (1989) suggested in the future research points (1, 2), the changing length of the standard box plot effects on the subjects' perception of the box length. As we found from the results of previous studies in the box plots. The same thing might happen in this study inaccuracy or biases in the subjects' judgments might occur with box length, box width and whisker length, as results of the visual illusion effect created by these interactions between the box length and, these two variables, box width and whisker length.

4- Method

Now will discuss the method used for these two experiments. The subjects were asked to make comparisons between box lengths of pair of vertical box plots placed side by side. Each box plot was in the center of an A4 sheet of paper, this applied to both the standard and the comparison plot. To give percentages for how much shorter, longer the length of the box plot (midspread) was than this in the booklet of the standard box plot.

Subjects: Subjects taking part in these two experiments were undergraduate third and fourth years from statistics department, Baghdad University, they were not familiar with the box plot, but had some knowledge of data analysis. There were (50) subjects taking part in box width experiment and (50) subjects were taking part in whisker length experiment; the subjects who had not understanding the instructions had their answers excluded from analysis.

Design: These experiments were building to examine the effect of box width and whisker length on the subjects' perception of the box plot length. The first experiment box width experiment, contains the forty five box plots, which were generated from level combinations of the two factors, box length with nine levels the middle level is the same as the box length of the standard plot, and box width with five levels one of the levels being similar to the box width of the standard plot. Each box plot was on an A4 sheet of paper and also a standard plot. And the second experiment whisker length experiment contains the forty five box plots, which were generated from levels combination of the two factors, box length with nine levels, one of the levels being similar to the box length of the standard plot. And whisker length with five levels, one of the levels being same to the whisker length of the standard plot. These length levels were determined by Cleveland & McGill (1984); who used the formula;

L_j = 10 x 10 (j-1) /12, (j=1 —-n), then we suggested 1 unit = 3 mm.

These values are equally spaced on a log scale and range from 10 to... N units, chosen values in order started by 10 units represented 9 box length levels, the middle one level of them represents the standard box plot length، which were different from the standard plots of Hussin, M.M. (1989,2006) there were the longest one and the smallest one. The nine levels represent the box length levels of the experimental box plots for these two experiments. These box length levels selected to fit the box plot on an A4 sheet paper, and present as large a range of plots as possible, with the other levels of variables, box width, and whisker length, sea variable levels in Table no.1A.

Table No.1A. levels of variables of these two experiments.

Length	Width	Whisker
L1 = 30	W1 = 10	S1 = 10
L2 = 36	W2 = 25	S2 = 30
L3 = 44	W3 = 40	S3 = 50
L4 = 54	W4 = 55	S4 = 70
L5 = 64	W5 = 70	S5 = 90
L6 = 78		
L7 = 94		
L8 = 114 L9 =140		

Standard box plot variables levels

L = 64 W = 40 S = 50

Materials: There were two booklets; each one contains forty five box plots. The first sheet in the booklet contained two examples of practice plot so that the subjects understood the experiment. Subjects were also given an instruction sheet, an answer sheet, and a standard box plot, the booklets were given to subjects in the lecture room, and each subject was given a booklet of one of the experiments. The instruction sheet asked subjects to compare the box plots from the booklet with the standard box plot. The subject was asked to give a percentage of how shorter or longer the length of the box plot was than that in the booklet of the standard box plot, and at all the times the standard box plot was the middle one without this being mentioned to the subjects. The subjects were also asked to write (T) or (B) respectively on the answer sheet if they thought that the length of the box plot on the booklet was longer than the standard box plot or vice versa. This provided a check on the direction of their judgments. The instructions asked subjects to make quick visual judgments rather than measurements. Examples of standard box plot, instruction sheet, answer sheet are not included, because the problem of the space.

٥ –*The statistical method used to analyze the data* .

The analysis of variance technique was used to analyze the data of these two experiments, these experiments were designed as repeated measures, and for such data the analysis of variance technique appears to be appropriate, for more detail see Hussin, M.M. (1989, 2006). The assumptions of the design can be summarized as:

$$X_{ij} \sim N(\mu_i, \sigma^2).$$

There were three models can be used for the analysis of variance technique fixed effects model, random effects model, and mixed effects.

The design of these two experiments were repeated measures design, and the model for this design is the special case from mixed mode1. In this design subjects are observed at all combinations of the independent variables, and the model for the first experiment the box width experiment is ;

$$Y_{ink} = U .... + W_1 + N_n + WN_{1n} + E_{1nk} .... (1)$$

'K' th observation (subjects).

'I' th level of box width factor (J) levels.

'n' th level of box length factor (I) levels.

In this model (1), the box length and box width are the fixed effects factors, and the subjects are a random effects factor. For this design as subjects are observed at all observations of the variables, it is expected that the observations on the same subjects will tend to be correlated, or be dependent. For this reason, this design needs more assumptions of homogeneity of the variance- covariance matrix.

1- The variances are :

$$\sigma^2_{x1} = \sigma^2_{x2} = \sigma^2_{x3} = \ldots = \sigma^2_{xn}.$$

2- The co variances are:

$$\sigma_{x1\,x2} = \sigma_{x1\,x3} = \sigma_{x2\,x3} = \ldots = \sigma_{xn-1\,xn}.$$

If this assumption is not met, it is impossible to use the usual F test, without some modifications. For this reason the conservative test provides approximation, but some times this test is negatively biased, (Winer, 1962, P. 306), for more detail see Hussin, M.M. (1989, 2006).

6- The box width experiment analysis of variance results.

This section will discuss the results of this experiment, and the fact that we choice the univariate analysis of variance, as the problems with the assumptions of normality as we found that some of data sets were light- tailed and some were skewed, or double peaked at upper and lower extremes, and also the violation of the variance- covariance matrix, we found the adjusted univariate analysis of variance with trend analysis is more suitable than the multivariate analysis.

This was recommended by Winer (1962, p. 306) by Rogan et al (1979, p. 269- 286), and by Huynh 1970, Huynh, 1978, Huynh et al, 1979, and also by Charles S. Davis (2002). The trend analysis also was found to study more specific aspects of the differences in patterns or shapes for the simple main effects of the variables in the analysis, and the polynomial contrasts is the best way to do this job. There are two variables in this experiments; box length with nine levels, and box width with five levels and the model is equation no.1 in the previous section.

Now let start with the results of the analysis of variance as in Table no.1, and started with the interaction effects of these two variables (WN_{1n}). It was found that the (F) value of this interaction effects (WN_{1n}) was equal to 11.293 and the tail probability for the usual (F) test was equal to (0.000). This means that the interaction effects have a high level of significance. But to use the usual (F) test for this design is highly restrictive because a design having correlated observations will affect the results in a positive bias in the usual (F) test. That is, the variance-covariance matrix should confirm the assumption of homogeneity of this test. Checking this assumption for the interaction by using the sphericity tail probability for the WN_{1n} revealed that the assumption of sphericity was not met, thus the conservative test provided an approximate test with the number of degrees of freedom for the F value reduced by (E) Epsilon. But even with this test whatever the reduce in the degrees of freedom as a large degree of heterogeneity in the variance – covariance matrix, still this test interaction effects have very high level of significance with this conservation test. Also it is very clear to recognize the interaction between these two variables from box width experiment plot no.1 and also can be seen from plot no.1 that the average of the absolute values of the errors decreases with the middle length levels as similar or near from the standard box length and increases with shorter and longer levels as far from the standard length.

And to examine the differences in trends of the variables' effects for the interaction, by using the polynomial contrasts, we found from Table no.1 twenty one polynomial components of this interaction significant. That means there are significant differences between all the trends of the interaction of these two variables, box length and box width as you can see from the Table no.1 and from plot no.1. These variables two of them are responsible for this interaction, and two of them are important for the subject' judgments or two of them have affected subjects' judgments. One possible explanation of these results is that as the box width changed, subjects judged area instead of length for area judgments and area is more difficult than length judgments, as the power law (Steven, 1975) suggested and Cleveland et al (١٩٨٤)found from their results. And also it might be as a result of the visual illusion effects on the subjects judgments created by the interaction between these two variables width / length as Cleveland et al (1987) accepted in their replies to the comments on their results.

This result of the interaction agrees with McCulloch's (1981), Knight's (١٩٨٢) results, and Hussin, M.M. (1989), group A experiment, box width experiments at Keele and Baghdad results, and Hussin, M.M. (2006). And disagree with Hussin, M.M. (١٩٨٩)، three factors experiment results, group B experiment at Baghdad. But this experiment result and Hussin, M.M. (2006) different from all above results as the interaction highly significant more than all the other experiments results and one explanation for these results are the standard box plot lengths in these experiments represent the middle level of the length levels and shorter than all the other box plots length levels of the experiment, which is different from all other experiments which agree with the suggestions of the further research points (1 ,2) to build these two experiments.

Now we consider the results of the main effect in Table no.1. It was found that the main effects of the two variables are significant. The F-value of the width variable equal to (2.863), and the tail probability for this variable is equal to (0.025), and that means high level of significance, and there is homogeneity of variance - covariance matrix. The sphericity assumption is met and the Mauchly's W test equal (0.943) and their significance level (0.973). We do not need to use the conservative test for this width variable. Now let us examine trend analysis for the width variable main effect, we found only the quadratic is significant with very high level of significance, and that can be seen from plot no.1 box width experiment.

In this experiment box width variable is important and had affected subjects' judgments of box length or the width variable misleads subjects in their judgments of box length, or we can put it down to visual illusion effects as created from the interaction between these two variables box width /length, but lease significance from other experiments by it self. This result agrees with Mc Culloch's (1981), Knight's (1982), and Hussin, M.M. (1989) results of the box width experiments for two groups A, and B experiments at Keele and Baghdad.

Consider the length variable in Table no.1. It was found that the F−value is equal to (18.453) and the tail probability for the F -value is equal to ٠.٠٠٠. The sphericity tail probability is (0.000). That would means this test has a very high level of significance with very large amount of heterogeneity of the variance - covariance matrix exists, the sphericity assumption is not met.

When conservative test, is used, the tail probability for F-value for this test is equal to 0.000. However, the F-value still has a very high level of significance.

This result disagrees with Mc Culloch's (1981) results, but agrees with Knight's (1982) result, and Hussin, M.M. (1989) results for all their experiments for the two groups of experiments of Keele and Baghdad Universities and also Hussin, M.M. (2006) results. Now let us consider the trend analysis. We can see that there are significant differences for all eight trends components the tail probability for the F - values for all of them are very high level of significance. We thought one possible explanation for these results, as we mentioned before; is the interaction between these two variables leads to these difficulties. The length variable very important for the experiment. Some of the subjects are excluded from the experiment because they did not follow the instruction.

٧- *The whisker length experiment analysis of variance results*.

In this section we will discuss the results of the analysis of variance with the trend analysis, the model for the experiment equation no. (2), two variables, box length with nine levels and whisker length with five levels.

$$Y_{n1k} = U... + N_n + S_1 + NS_{n1} + E_{n1k} ..... (2)$$

' k ' the observation (subjects)

' I ' the level of whisker length factor (J) levels .

' n ' the level of box length factor (I) levels.

This model repeated measures design, the box length and whisker length are the fixed effects factors, and the subjects are a random effects factor. Now let us consider Table no.2, and begin with the interaction effect of these two variables box length and whisker length. It was found that the F value for this test was equal to 9.343, and the tail probability for the F – value was equal to 0.000. This means that this interaction is significant with very high level of significance, and the sphericity tail probability for the interaction is equal to 0.000. The sphericity assumption is not met. It was for this reason that the conservation test was used, and still the tail probability for the F - value of this test is equal to 0.000, and we do not need to use the Greenhouse & Geisser, because the test very highly significant. Therefore, this means there is a very high level of significance for the interaction effects. Let us now examine the trend analysis of this interaction, we found twenty two trend components were significant in table no.2, with a very high levels of significance and also can see that very clear from plot no.2.

This means that the (NS) interaction arises from the differences between all components of trends for these two variables.

This results of the interaction different from the Knight's (1982) result, and different from Hussin, M.M. (1989) result, whisker length experiment, group A at Baghdad, and also different from Hussin, M.M. (1989) result of the whisker length experiment, group B at Keele and Baghdad. But agrees with Hussin, M.M. (١٩٨٩) results of the whisker length experiment, group A at Keele, and with Hussin, M.M. (٢٠٠٦) results of the whisker length experiment, And we thought this results different from the results of Hussin, M.M. (1989) whisker length experiments group B, at Keele and Baghdad , because the length of the standard box plot is the middle length levels of the box plot lengths in the booklet of this experiment which it was the standard box length in the whisker length experiments, group B, of the Hussin, M.M. (1989) was longer than all the other in the booklet of the experiments. In this experiment the subjects faced problems in their judgments, the reason might be that the joining of these two variables creates an interaction or perceptual problem as Lovie (1985) argued, and the visual illusion might then affect the subjects' judgments as Cleveland et al (1987) accepted.

Now let us consider the results of the main effects of these two variables in Table no.2. It was found that the F-value of the box length is equal to 60.252, and the tail probability of this test is equal to 0.000. The sphericity tail probability of the F-value is equal to 0.000, the assumption of the variance - covariance matrix is not met. The conservation test should be used, and still the F-value has a very high level of significance higher than all the other in previous work in this area of box plot except Hussin, M.M. (2006) results the same. One possible explanation of this result is that the box length of the standard box plot is the middle level lengths of the other box lengths of the booklet, and also can be seen very clear from plot no.2 that the average of the absolute values of the errors decreases with the middle length levels and increases with shorter and longer levels. Now let us examine the trend test of this effect. It was found that five trend components are significant with very high levels of significance, this result can be seen from plot no. 2 and also multivariate Tests Table no. 2, and these components are responsible for the high significance of box length main effect. This variable very important for the experiment, and Wber's Law might help to explain the results.

Finally let us examine the whisker length main effect, it was found that the F-value was equal to 8.392 in Table no.2, and the tail probability for the F-value was equal to 0.000. The sphericity tail probability of the F-value is equal to 0.107. Now we don't need to use the conservative test because the assumption of the variance - covariance matrix is met. Now let us examine the trend test of this main effect, it was found that the

Linear, and the Cubic and order 4 trends are significant with very high level of significance equal to 0.004 or higher, and these components are responsible for the significance of the whisker length main effect, and this result can be seen from plot no.2 also. This variable whisker length very important for this experiment and affected subjects' judgments of box length, even when the subjects were asked to concentrate on box length, but not in obvious pattern as with length variable. This result agree with Hussin, M.M. (1989) result, of whisker length experiments group A, at Keele, and Baghdad, but not in the same level of significance as in this experiment very higher than all the others, and agree with Hussin, M.M. (2006) results with similar level of significance. And this result is disagree with Knight's (1982) results, and also disagree with Hussin, M.M.'s (1989) results of whisker length experiments, group B, at Keele and Baghdad. One possible explanation of this result it might be that the different box length of the standard of this experiment from the other experiments.

٨- *Conclusion*

The idea behinds the constructions of these two experiments, we suggested that the variations on a basic box plot are important and affected subjects' judgments. And changing the standard box plot length influence a subjects' perception of box length, and also changing the length levels of the box plots from the experiments of Hussin, M.M. (1989) group B. And we found that from the results of these two experiments.

The conclusions of these two experiments can be summarized as follows:

1. **The standard box plot length very important and influence subjects' judgments as we found from this experiment results and the same results of Hussin, M.M. (2006) specially whisker length experiment, and the results agrees with Hussin, M.M. (1989) suggestion in future research points(1, 2).**
2. **The variables interactions of the box plot are very important and impaired subjects' judgments of box length. And that might be as a visual illusion effects created by these interactions, or as Lovie (1985) put it perceptual problems with judging box plot.**
3. **The whisker length variable is a very important variable and influences a subjects' judgments of the box length, even this variable irrelevant variable in the experiment. But still lease than the length variable. And this result and, Hussin, M. M. (2006) result for the same experiment, different from all previous experiments results of the whisker length variable of the box plot, as it is important in this experiment.**

4- The box plot length (midspread) was the most important variable to affect subjects' judgments in the box plot, because the box length variable the relevant variable in the experiment so this result very reasonable, these results agree with Knight's (1982) results, and with Hussin, M. M.'s (1989,2006) results for all box plot experiments.

5- The box plot width is a very important variable in the experiment, but it might be lease than to the box length and that's very fair and reasonable, but the interaction between these two variables make the subjects' judgments more difficult, as we mentioned before it might be create visual illusion.

6-Tese variations on a basic box plot very important because add more information to the box plot, but the coast more difficulties arises.

7- Some of the subjects face some difficulties in making the box plot judgments, so this agree with Lovie (1985) put it perceptual problems. with judging box plot.

BIBLIOGRAPHY

١ - Baldwin, J.M. (1895). The effect of size contrast upon judgments of position in the retinal field, Psychological Review, 2, p.244-59. Cited in Robinson, 1972.

٢- Chambers, J.M., Cleveland, W.S., Kleiner, B. and Tukey, D.A. (1983). Graphical methods for data analysis, Duxbury (softcover), Boston and Wadsworth (hard cover), Belmont, California.

3- Charles S. Davis (2002).Statistical Methods for the Analysis of Repeated Measurements. Springer-Verlag New York, Inc.

٤ - Cleveland, W.S., Harris, C.S. and McGill, R. (1982). Judgments of Circle Size on Statistical Maps, Journal of the Amarican Statistical Association 77 p.541-547

٥ -Cleveland, W.S. & McGill, R. (1984). Graphical perception: Theory, Experim- entation, and application to the development of graphical methods. Journal of the American Statistical Association, 79, p.531-554.

٦ -Cleveland, W.S. & McGill, R. (1986). An experiment in graphical perception. International Journal of Machine Studies, 25, p.491-500.

٧ - Cleveland, W.S. & McGill, R. (1987). Graphical perception: The visual decoding of quantitative information on graphical displays of data,
(with discussion). Journal of the Royal Statistical Society A, 150, part3, p. 192-229.

٨ - **Greenhouse, S.W. and Geisser, S. (1959). On methods in the analysis of profile data. Psychometrika, 24, p.95-112**.

٩ - **Hussin, M. M. (1989). Some studies of a graphical method in statistical data analysis; subjective judgments in the interpretation of Box plots– Unpublished Ph.D. Thesis, Keele university , U.K**. .

10- Hussin, M. M. (2005), Statistical Data Analysis Using SPSS.

11- Hussin, M. M. (2006), An Experiments on the Box plot, Journal of Administration and Economy, Volume 12, number 42 .

١٢ - **Huynh, H. (1978). Some approximate test for repeated measurements designs. Psychometrika, Vol.43, No2, p. 161-175**.

13 - Huynh, H. and Fedit, L. (1970). Conditions under which mean square ration in repeated measurements designs have exact F-distributions. Journal of the American Statistical Association, Vol. 65, No.332, p. 1582-1589.

١4 - **Huynh, H. and Mandeville, G.K. (1979). Validity conditions in repeated measures designs. Psychological Bulletin, Vol.86, No.5, p.964 -973**.

١5- **Knight, L.C. (1982). "Do additions to the boxplot confuse the viewer "؟ Unpublished BSc dissertation - University of Keele**.

١٦ - **Lovie, P.(1985). Subjective judgment in data analysis. Research Report, Department of Mathematics, University of Keele. Alsopresented in International Seminar on Cognitive processes in Mathematics, Keele, 1985**.

١٧ - **Lovie, A.D. & Lovie, P. (1987). Discussion of the paper by Drs Cleveland & Mc Gill. J.R.Statist. Soc. A., p.210-229**.

١٨ - **Loynes, R.M. (1987), Discussion of the paper by Drs Cleveland &McGill, J.R.Statist. Soc. A., p.210-229** .

١٩ - **Mc Culloch, J. (1981). The variable- width box plot from different perspective. Unpublished B.A. dissertation, University of Liverpool**.

٢٠ - **McGill, R., Tukey, J.W. and Larsen, W.A. (1978). "Variations of box plots", The American Statistican, 32, p. 12-16**.

٢١ - **Rogan, J.C., Keselman, H.J. and Mendoza, J.L. (1979). Analysis of repeated measurements. British Journal of Mathematical and statistical Psychology, 32, p.269-286**

٢٢ - **Simkin, D. and Hastie, R. (1987). An information processing analysis**

of graph perception. Journal of the American Statistical Association, Vol.82, No.398, p.454-465.

٢٣- Sim, C.H. ;Gan, F.F ;Chang, T.C.(2005). Outlier Labeling with Box plot Procedures. Journal of the American Statistical Association ,Volume 100, Number 470, June ,PP. 642-652.

٢٤ - Stevens, S.S. (1975). Psychophysics. Wiley, New York.

٢٥- Titchener, E.B. (1906). Experimental psychology. New York, Macmillan. Cited in Murch, G.M. (1973). Visual and Auditory perception, New York, The Bobbs-Mervill Company, Inc.

٢٦ - Tukey, J.W. (1977). Exploratory data analysis. Reading mass: addision– Wesley.

٢٧ - Winer, B.J.(1962). Statistical Principles in Experimental Design. New York,Mc Graw-Hill .