CHAPTER V

SOFTWARE EVALUATION

Three software evaluation methods, Questionnaire, Statistic and Meta-analysis were described respectively in this chapter.

5.1 Questionnaire

The questionnaire was developed to evaluate student's attitude on this EM-CAI. It has been shown in Appendix III

5.2 Statistical Evaluation

Evaluation by ststistical method was performed stepwise as following [13].

5.2.1 Sampling the population

The sampled population are obtained by using the three different major students, the faculty of Engineering, Mahidol University that must attend the Engineering Metallurgy Course. Sample are classified into two groups, control group and treatment group (CAI group). The control group learns by using traditional instruction and CAI group learns by supplemented with the EM-CAI.

The sample size in each groups are about 50. This number are considered large enough to be the representative sample and have a normal distribution.

5.2.2 Collect data as test score from the two target groups.

Control group and CAI group are tested in the same conditions. The scores are recorded separately to process in further step.

5.2.3 Hypothesis testing

5.2.3.1 Define the null hypothesis (H_0) and the alternative hypothesis (H_i) using the research problem criteria.

The research problem is to evaluate whether learning with supplement of CAI is more effective than traditional learning. Therefore a directional hypothesis is used and one-tailed test is performed using the mean test score for statistical testing. The null and alternative hypothesis (H_o and H_i) are defined as following:

$H_o : \overline{x}_1 - \overline{x}_2 \leq 0$

(test score of CAI group, x_1, is equal to control group, x_2)

$H_i : \overline{x}_1 - \overline{x}_2 > 0$

(test score of CAI group, x_1, is higher than control group, x_2)

Where $\overline{x}_1$ = mean test score of CAI group.

$\overline{x}_2$ = mean test score of control group.

If the statistical test result refuse the null hypothesis, the alternative hypothesis is accepted. Thus, it can be concluded that mean test score of CAI group is higher than that of control group.

5.2.3.2 define the significant level (α)

The significant level at 0.05 is selected.

5.2.3.3 Calculate the test statistic (t-test)

Using the collected data to calculate the statistic. In this thesis, t-test is used for testing. The mean test score from the two sampled groups are compared using t-test (two independent samples). The statistical parameters employed for testing include of Mean ($\overline{X}$) and Standard deviation(s). The calculation formula is :

CASE 1: if variance of two independent samples are equal

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}}$$

when

$$df = n_1 + n_2 - 2$$

$$s_p^2 = \frac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2}{n_1 + n_2 - 2}$$

when

s_p^2 = total variance of the two groups

$\overline{X}_1$ = mean test score of the CAI group

$\overline{X}_2$ = mean test score of the Normal group

s_1^2 = sample variance of the CAI group

s_2^2 = sample variance of the Normal group

n_1 = number of samples in the CAI group

n_2 = number of samples in the Normal group

CASE 2: if variance of two independent samples are not equal

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$\text{and} \quad df = \frac{[(s_1^2/ n_1) + (s_2^2/ n_2)]^2}{\frac{(s_1^2/ n_1)^2}{n_1-1} + \frac{(s_2^2/ n_2)^2}{n_2-1}}$$

Note: all parameters are the same definition as in 5.2.3.3.

5.2.3.4 Finding critical value

Critical value is used as a decision criteria. The value is determined from region of retention. In one-tailed test, the region of rejection is equal to a significant level. For example, at significant level 0.05 of positive direction one-tailed test, will be determined from 1.0-0.05 = 0.95. Use 0.95 to find the area under normal distribution curve of t- value. The t-value obtained is compared with t-value from the t-distribution table.

5.2.3.5 Making statistical decision

Statistical decision employs the calculated t-value. If calculated t-value $\geq$ t-value from table, then refuse the null hypothesis and accept the alternative hypothesis.

5.2.3.6 Conclusion

Conclusion is makes by using the result of statistical decision. In this thesis, for example, if the decision result find that the null hypothesis is refused, the conclusion would be "the mean test score from the CAI group is higher than the control group, or learning with supplemented of CAI is more effective than traditional learning only".

5.3 Meta-analysis Method

By using t-test, when the null hypothesis is rejected, at most the conclusion is that "the difference is not zero". This conclusion can not specify about how large the difference one should expect. Figure 5.1 shows that difference with arrows.

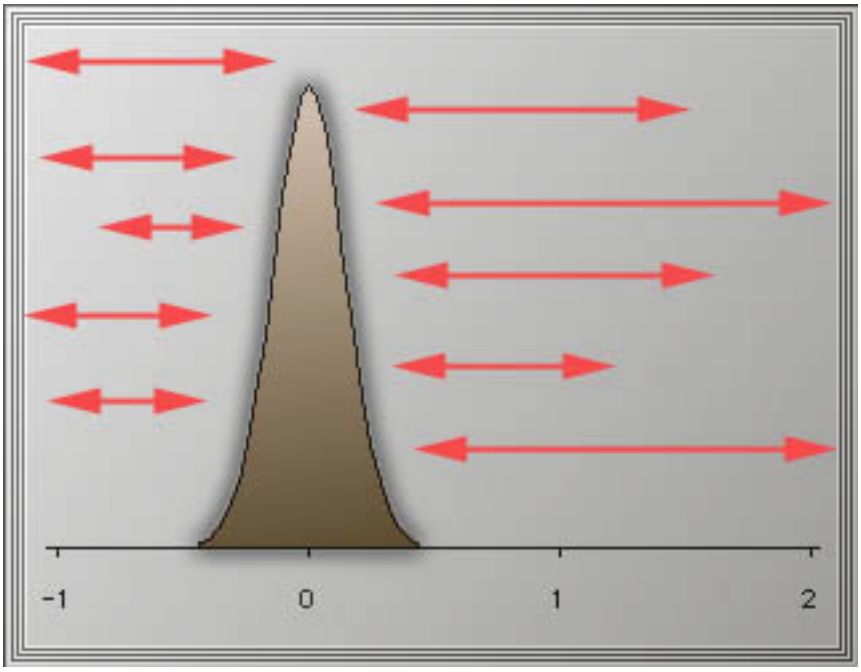

Figure 5.1 The length of arrows indicate the possible difference value of control group and sample group.

This problem can be solved by using Meta-analysis method [5,14] as stated in section 2.3. In this method, Effect size value (Es) was calculated and compared to the conventional values of Effect size as shown in Table 5.1. Effect size is the mean difference between group in the form of standard score, i.e. the ratio of the difference between the mean to the standard deviation as shown in equation:

$$\text{Effect size} = \frac{|\text{Mean of control group} - \text{Mean of treatment group}|}{\text{Standard deviation of the control group}}$$

This value shows the minimum difference that worth in research attention. The larger the Effect size (the difference between the null and alternative means) is, the greater the power of a test is.

Table 5.1 The conventional values of Effect size in different level.

Level of Effect size	Value range of Effect size (d)
Small	0.20 - 0.49
Medium	0.50 – 0.79
Large	≥ 0.80

Comparing calculated Effect size to the conventional Effect size make researcher conclude that whether the difference of null hypothesis and alternative hypothesis is worth in research attention. Figure 5.2 shows the considered Effect size, indicated by arrow.

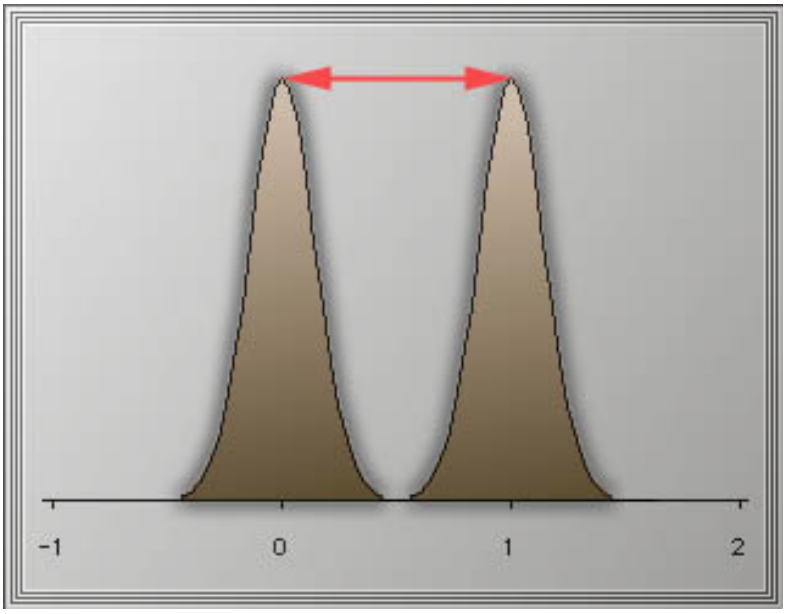

Figure 5.2 The arrow shows the mean difference between control group and sample group

In this research, the EM-CAI is also evaluated by the Meta-analysis. Table 5.1 will be used as the benchmark to judge the effectiveness of EM-CAI.