

This study was to compare the reliability and validity of a multiple-choice test from different scoring techniques: the conventional and the confidence weighting techniques; in measuring the two groups of different Cognitive Domain levels which were the lower levels (the first three levels) and the higher levels (the last three levels).

The results of this study were:- The reliability and validity in the high, moderate and low academic achievement groups. 1) By using the conventional scoring technique to the lower levels test, the reliability of the test between the high and moderate or low groups were significant difference ($p < .01$); whereas there were no significant difference between the high and moderate group. The validities of the test in all groups were not significant difference. 2) By using the conventional scoring technique to the higher levels test, the reliabilities and validities of the test in all groups were not significant difference. 3) By using the confidence weighting scoring technique to the lower levels test, the reliabilities of the test in all groups were not significant difference. The validity of the test between the high and low groups was significant difference ($p < .05$); whereas there were no significant difference between the moderate and the high or low groups. 4) By using the confidence weighting technique to the higher levels test, the reliability of the test between the high and low groups was significant difference ($p < .05$); whereas there were no significant difference between the moderate and high or low groups. The validities of the test in all groups were not significant difference.

:- The reliability and validity of the test which scored by two different techniques: the conventional and the confidence weighting techniques. 1) There were no significant difference between the reliabilities of the lower levels test scoring by two different techniques in the high, moderate and low groups. But the validity of the test scoring by two different techniques was significant difference in the high group ($p < .01$); whereas there were no significant difference in the other two groups. 2) There were no significant difference between the reliabilities of the higher levels test scoring by two different techniques in all groups. But the validities of the test scoring by two different techniques was significant difference in the low group ($p < .01$); whereas there were no significant difference in the other two groups.