

EXHIBIT 7.5 Framework for Reliability of Measures

Person	Scores on Attribute A							
	Objective (Test Items)				Subjective (Raters)			
	Time 1		Time 2		Time 1		Time 2	
	X_1	Y_1	X_2	Y_2	X_1	Y_1	X_2	Y_2
A	5	5	4	5	5	5	4	5
B	5	4	4	3	5	4	4	3
C	5	5	5	4	5	5	5	4
D	5	4	5	5	5	4	5	5
E	4	5	3	4	4	5	3	4
F	4	4	4	3	4	4	4	3
G	4	4	3	4	4	4	3	4
H	4	3	4	3	4	3	4	3
I	3	4	3	4	3	4	3	4
J	3	3	5	3	3	3	5	3
K	3	3	2	3	3	3	2	3
L	3	2	4	2	3	2	4	2
M	2	3	4	3	2	3	4	3
N	2	2	1	2	2	2	1	2
O	1	2	3	2	1	2	3	2

NOTE: X_1 and X_2 are the same test item or rater; Y_1 and Y_2 are the same test item or rater. The subscript "1" refers to T_1 and the subscript "2" refers to T_2 .

internal consistency reliability. The relevant comparisons are X_1 and Y_1 , and X_2 and Y_2 . It is hoped that the comparisons will show high similarity, because both items are intended to measure A within the same time period.

Now treat the four sets of scores as coming from the subjective measure, which relied on raters. Comparisons of these scores involve what is called interrater reliability. The relevant comparisons are the same as with the objective measure scores, namely, X_1 and Y_1 , and X_2 and Y_2 . Again, it is hoped that there will be high agreement between the raters, because they are focusing on a single attribute at a single moment in time.

Comparisons Between T_1 and T_2

Comparisons of scores between time periods involve assessment of measurement stability. When scores from an objective measure are used, this is referred to as test-retest reliability. The relevant comparisons are X_1 and X_2 , and Y_1 and Y_2 . To the extent that A is not expected to change between T_1 and T_2 , there should be high test-retest reliability.