Correlation Paradox

istation Whitepaper 2009-01

Rev A

September 2009

800 E Campbell Rd, Ste 224
Richardson, TX 75081
(866) 883-READ
info@istation.com

Summary

In research, when correlations from two or more groups are combined, a paradoxically different pooled correlation is sometimes observed. This phenomenon, commonly referred to as Simpson's Paradox, is explained through the analysis of heuristic data.

Correspondence concerning this whitepaper should be addressed to Dr. Kevin E. Kalinowski, Director of Research, istation, 800 East Campbell Road, Suite 224, Richardson, Texas 75081. E-mail: KKalinowski@istation.com

Correlation Paradox

When reading correlation tables in a validity report, you might wonder how several correlations for individual grades combine to form a higher pooled correlation. For example, you might observe the fictitious results in Table A and wonder how the three moderate .447 correlations combine to create a nearly perfect correlation of .964.

Table A
Correlations between an External Measure and ISIP for grades K–2

K	*1*	*2*	*K-2*
.447	.447	.447	.964

Let's look at the data that produced these results. For this example, five students from each grade took both ISIP and some external measure (ExtMeas). Their data are shown in Table B.

Table B
Heuristic Data

Student	*Grade*	*ISIP Score*	*ExtMeas Score*
A	K	160	5
B	K	165	2
C	K	170	2
D	K	175	4
E	K	180	7
F	1	190	15
G	1	195	12
H	1	200	12
I	1	205	14
J	1	210	17
K	2	220	25
L	2	225	22
M	2	230	22
N	2	235	24
O	2	240	27

Table B shows that ISIP scores increased from 160 to 240 with no repetition in scores. ExtMeas scores generally increased across the fifteen students, but not monotonically as with the ISIP scores.

Upon closer inspection, it can be seen that the ExtMeas scores for the five students in Kindergarten are between 2–7, the First grade scores are between 12–17, and the Second grade scores are between 22–27. In fact, 10 points were simply added to the Kindergarten ExtMeas scores to get First grade ExtMeas scores, and 10 more points were added to get Second grade ExtMeas scores. Similarly, 30 and 60 points were added to the Kindergarten ISIP scores to obtain First and Second grade ISIP scores, respectively.

If only the Kindergarten scores are plotted, as in Figure 1, it is fairly clear why their correlation is moderate.

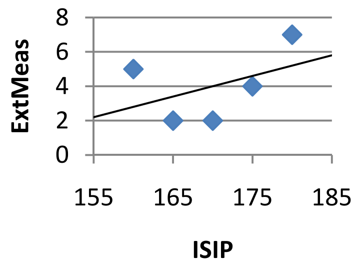

Figure 1. Scatterplot of ISIP and ExtMeas for the 5 Kindergarten students.

A general increasing trend is seen in the data, but the overall effect is clearly "u" shaped. Additionally, the plotted line in Figure 1 is a linear trend line, and can be thought of as the best representation all of the data using a straight line. Because the data are nonlinear in nature, the straight line does not do a very good job of capturing the five data points. Similarly, if only First grade or Second grade scores were plotted, the shape of their graph would be identical to Figure 1 but the numbers would be different. For each of the three groups, the correlation between the ISIP and ExtMeas scores is .447.

Now let's plot all of the points, Kindergarten through Second grade, as in Figure 2.

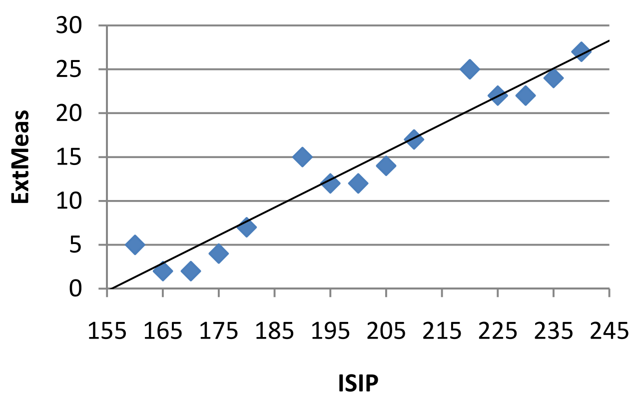

Figure 2. Scatterplot of ISIP and ExtMeas for all 15 students.

Notice that the three distinct "u" shapes, one for each grade, are still present. But when all of the data are taken together, the overall linear (straight line) trend overwhelms the 3 nonlinear (u-shaped) trends, and the plotted linear trend line does a good job of capturing most of the data points. As a result, we get a nearly perfect correlation of .964 for the entire K–2 sample.

Although not discovered by the British statistician, Edward Simpson, this general concept was described by Simpson (1951), and is commonly referred to as Simpson's Paradox. The Paradox is often used in statistics education to illustrate the care statisticians need to take when interpreting data.

References

Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society, 13*, 238–241.