

The paper proposes an index for detecting cheating. Although the title uses the term “collusion,” the given index does not go so far to detect this more specific case of cheating. *Collusion* implies the awareness of the act by at least two parties, namely the source and the suspected cheater. Although the proposed index is a more generalized one, it cannot possible detect the participation of the source.

Existing indices use the number of matching items that both the source and suspected cheater answered wrong. The proposed index tries to improve on this by also considering the number of matching items that both source and suspected cheater answered right.

All indices presented (whether based on classical test theory or item response theory) use the tail probability of some distribution as the basis of their measures. The existing K indices use the binomial distribution (they differ only on the way the expected number of wrong answers of the source, is computed). The proposed index uses the trinomial distribution and tries to be more flexible in modeling by considering the possibility of a blank answer.

Some improvements can be suggested. One, the parameters can be subjected to item response theory. By modeling the probability of the number of wrong answers as a function of item difficulty and the source’s ability, a more realistic model can be arrived at. This is more in line with reality since if a particular item has a high degree of difficulty then a student might refuse to answer it.

Another suggestion has to do with simulation. It is suggested that the number of items greatly exceeds the number of students. This is so because of the possibility of having a small number of students belonging to the set of students having the same number of wrong answers as the suspected cheater.

A final suggestion is to study the possibility of using the “number of students with the same number of right answers as the cheater” as the denominator m^R since this number can vary (albeit not too much) from n_c .

Also, its suggested to drop “*and its Statistical Properties*” from the title since this was not thoroughly discussed in the paper.