Commentary

Comment on Faulder Colby "Using the binomial distribution to assess effort: Forced-choice testing in neuropsychological settings"

Richard I. Frederick
US Medical Center for Federal Prisoners, 1900 W. Sunshine St., Springfield, MO 65807, USA
E-mail: rfrederick@bop.gov

Much of what Colby writes is unnecessary and some of it is definitely muddled. Colby is mistaken about probabilities of correct (p) and incorrect (q) responses in a two-alternative forced-choice (2AFC) procedure when the answers are equally tenable because of the absence of any ability to discern between them; in such circumstances it is always correct to assume $p = q = 0.5$. Of course, one can identify innumerable binomial (two-outcome) distributions in which $p \neq q$, for example, correctly guessing a number between 1 and 100, $p = 0.01$, or guessing incorrectly, $q = 0.99$. To be sure, when p and q represent the equivalent likelihoods of heads or tails, bending a coin changes p and q – a matter which does not contradict the equivalence of p and q when the coin is fair.

Constructing 2AFC paradigms to generate chance distributions of right and wrong responses from naive individuals (i.e., "null hypothesis" procedures; $p = q$) to identify instances in which the observed p is strikingly lower than the p expected by chance has proven a generally unfruitful means of identifying malingering. Adopting a different strategy with a 2AFC paradigm that ignores p and q and focuses on distinctions in patterns of responding among truly impaired individuals, truly normal individuals, and individuals feigning impairment has generally proven much more fruitful. This is what Tombaugh [4] has accomplished with the TOMM. We can refer to this as testing an "alternate hypothesis".

Colby has articulated a clever twist on the "alternate hypothesis" strategy. One can establish different values of expected minimal p and expected maximal q by researching the error rates of truly impaired individuals. That is, it is possible to identify situations in which the probability of choosing correctly between two answers is greater than 0.50 (e.g., see [2]). Generating expected p values based on information about how impaired individuals perform should prove much more powerful than routinely accessing a "chance" distribution to generate the expected p value. This paradigm has been reported elsewhere as the "floor effect" strategy [1,3].

I suspect most people who do poorly on the TOMM do so because they intend to do poorly. I suspect that most people who intend to do well on the TOMM, but who exert less than maximal effort, still do quite well on the TOMM. In other words, I do not see the TOMM as an "effort test". I encourage Colby and other authors to clarify distinctions between "intention" and "effort" and to investigate which methodologies best capture these separate elements of feigning.

References

[1] R.I. Frederick, A personal floor effect strategy to evaluate the validity of performance on memory tests, *Journal of Clinical and Experimental Neuropsychology* **22** (2000), 720–730.

[2] R.I. Frederick and R.L. Denney, Minding your 'p's' and 'q's' when conducting forced-choice recognition tests, *The Clinical Neuropsychologist* **12** (1998), 193–205.

[3] R. Rogers, E.H. Harrell and C.D. Liff, Feigning neuropsychological impairment: A critical review of methodological and clinical considerations, *Clinical Psychology Review* **13** (1993), 255–274.

[4] T.N. Tombaugh, *TOMM: Test of Memory Malingering manual*, Multi-Health Systems, Toronto, 1997.