

# Using Potential Performance Theory to Assess Differences in Math Abilities between Citizens from India and the United States

Stephen Rice<sup>1</sup> & David Trafimow<sup>1</sup>

<sup>1</sup> Potential Performance Technologies, LLC, Las Cruces, New Mexico, USA

Correspondence: Stephen Rice, Potential Performance Technologies, LLC, Las Cruces, New Mexico, USA. Tel: 1-575-571-7750. E-mail: [scrice67@gmail.com](mailto:scrice67@gmail.com)

Received: June 21, 2012 Accepted: July 9, 2012 Online Published: July 18, 2012

doi:10.5539/hes.v2n3p24

URL: <http://dx.doi.org/10.5539/hes.v2n3p24>

## Abstract

For years, researchers and academics have known that American students perform more poorly on average compared to students from some other countries, including India. The usual explanation is that some systematic factor (e.g. knowledge, skill set, test-taking ability, etc.) is responsible for the differences. The current study examines the issue from a different perspective; we assess the consistency in which participants performed in an Algebra test, and used this consistency to determine their potential performance. Participants were randomly selected from India and the United States and were given a 50-question algebra test, followed by a break, and then followed by the same test again. The data revealed that while the Indian participants scored about 8% higher on the test, the majority of their performance increase was due to being more consistent than their American counterparts.

**Keywords:** PPT, potential, performance, theory, mathematics, India, United States

## 1. Introduction

It is well known that an emphasis on mathematics education, and taking more mathematics classes, is associated with better performance on problems in mathematics. In countries where there is more focus on education in mathematics, students tend to have superior mathematics performance compared to students in countries with less of a focus on mathematics education (e.g. Phelps, 2001). Although the association of mathematics experience with performance in mathematics is indisputable, the underlying reason for this association is not completely clear. Our goal is to attempt a first investigation of this general issue, using algebra as the subject of interest. However, to understand the rationale for our investigation, it is important to distinguish between *consistency* and *potential performance* as possible explanations.

### 1.1 Potential Performance Theory (PPT)

The distinction between consistency and potential performance comes out of potential performance theory (hereafter, PPT; Trafimow & Rice, 2008), and has been supported by numerous empirical studies as well as simulations (Hunt, Rice, Trafimow & Sandry, in press; Rice, Geels, Hackett, Trafimow, McCarley, Schwark, & Hunt, 2012; Rice, Geels, Trafimow & Hackett, 2011; Rice & Trafimow, in press; Rice, Trafimow & Hunt, 2010; Rice, Trafimow, Keller, Hunt & Geels, 2011; Trafimow, Hunt, Rice & Geels, 2011; Trafimow, MacDonald & Rice, in press; Trafimow & Rice, 2008; 2009; 2011). The idea is that there are two general classes of impediments to performance on any task, including algebra problems. One class of impediments constitutes systematic factors (e.g., simply not knowing the material) and another class constitutes factors leading to a lack of consistency (e.g., a door slams in another room that provides a distraction while choosing the answer to a problem). Appendix A provides an abstract overview of how to conduct a PPT study, but we wish to further explain it here with the specific example of taking an algebra test.

Let us consider consistency first, as it is the most important PPT concept to understand. In the usual PPT paradigm, the researcher has participants complete two blocks of trials rather than one block, and every problem on the first block is repeated on the second block. Given that every participant completes two blocks of trials, it is possible to compute a correlation coefficient across the two blocks of trials, for each and every participant. This correlation coefficient measures each participant's consistency across the two blocks of trials and it also can be called a "consistency coefficient." We emphasize that consistency does not refer to doing the same thing on

every trial (e.g., choosing “true” on a true-false test), but rather two making similar responses on similar trials across both blocks of trials.

As has been known since at least Spearman’s (1904) seminal work, it is a fact of statistical regression that inconsistency pushes observed scores towards the chance level. An easy way to see this is to imagine a person who knows extremely well how to work all of the algebra problems on a test, but on five of the problems, her actual answer is the result of a coin flip. Or suppose a coin flip is used for 10 of the problems, or 15, and so on. There are two consequences. First, consistency coefficients will decrease. Second, actual performance will decrease. As Spearman showed, these effects are related. On average, as consistency decreases so will observed performance.

Let us now consider observed performance. Suppose that people complete 50 true-false algebra problems, take a break, and then complete them again. There are four possible combinations of correct answers and actual ones: (a) the person answers “true” and the correct answer is “true,” (b) the person answers “true” and the correct answer is “false,” (c) the person answers “false” and the correct answer is “true,” and (d) the person answers “false” and the correct answer is “false.” Because there are many items, rather than only one item, it should be easy to see that a person could have any particular frequency of each of these four combinations. In addition, and this is crucial for PPT, it is possible to convert these cell frequencies into a correlation coefficient, using Equation 1 below, where  $r$  is the observed correlation coefficient and  $a$ ,  $b$ ,  $c$ , and  $d$  are the observed frequencies for each of the four possible combinations. Because the row 1 frequency is the sum of cells  $a$  and  $b$ , the row 2 frequency is the sum of cells  $c$  and  $d$ , the column 1 frequency is the sum of cells  $a$  and  $c$ , and the column 2 frequency is the sum of cells  $b$  and  $d$ , it is possible to use column frequencies ( $R_1$ ,  $R_2$ ,  $C_1$ , and  $C_2$ ) rather than cell frequencies in the denominator of Equation 1.

$$r = \frac{|ad-bd|}{\sqrt{(a+b)(a+c)(b+d)(c+d)}} = \frac{|ad-bd|}{\sqrt{R_1 R_2 C_1 C_2}} \quad (1)$$

Based on Spearman’s (1904) work, and assuming that two blocks of trials were used and a consistency coefficient computed, it is a simple matter to adjust the correlation coefficient in Equation 1 for the person’s consistency coefficient. In Equation 2, we let  $R$  denote the performance correlation that would have been obtained had the person been perfectly consistent. Another way to think about  $R$  is that it denotes the correlation that would correspond to the person’s *potential performance*, in the absence of randomness. We will return to this notion shortly. In addition, we let  $r_{XX'}$  denote the consistency coefficient.

$$R = \frac{r}{\sqrt{r_{XX'}}} \quad (2)$$

Then, if  $R$  denotes the correlation coefficient that would correspond to potential performance that would be obtained if the person’s performance were perfectly consistent, how do we actually obtain an estimate of the person’s potential performance? To answer this question, it is first necessary to obtain potential cell frequencies, which are the frequencies in each of the four combinations we discussed earlier, but under the condition of perfect consistency; these will be denoted by upper case letters  $A$ ,  $B$ ,  $C$ , and  $D$ , respectively, to distinguish them from  $a$ ,  $b$ ,  $c$ , and  $d$  as we used for actually observed cell frequencies.

$$A = \frac{R\sqrt{R_1 R_2 C_1 C_2} + C_1 R_1}{(R_1 + R_2)} \quad (3)$$

$$B = R_1 - A \quad (4)$$

$$C = C_1 - A \quad (5)$$

$$D = R_2 - C \quad (6)$$

Given that one has obtained  $A$ ,  $B$ ,  $C$ , and  $D$  for a person, that person’s *potential performance* or *potential score* is easy to compute. Equation 7, below, provides the equation.

$$\text{potential performance} = \text{potential score} = \frac{A+D}{A+B+C+D} \quad (7)$$

Let us summarize thus far. Every person who takes an algebra test has an observed score, which is the proportion correct. In addition, however, given two blocks of trials, it is possible to compute a consistency coefficient. And using both the observed score and the consistency coefficient, it is possible to estimate a potential score, which renders the person’s performance under the context of perfect consistency. Because a lack of perfect consistency causes observed scores to be closer to chance than are potential scores, potential scores are generally larger than observed scores. More generally, PPT renders it possible to parse the reasons for a lack of perfect observed

performance into the deleterious effects caused by random factors (inconsistency) or systematic ones (low potential score).

Given the foregoing, it is now possible to address our main goal. Because lack of consistency (that reflects deleterious random effects on observed performance), as well as low potential performance (that reflects deleterious systematic effects on performance) are potential candidates for imperfect algebra performance, we might ask why people in at least some other countries tend to outperform Americans. One possibility is that the difference is due mostly to a difference in potential scores (i.e. systematic factors). The other possibility is that the difference is due mostly to a difference in consistency. Our goal is to investigate this issue with samples from the US and India.

## 2. Experiment

In the experiment, participants were recruited from either the United States or from India, using an online survey system in order to collect the relevant data.

### 2.1 Participants

Eighty participants from the online community participated in the experiment. Forty (22 females) participants were located in the United States with a mean age of 32.38 years ( $SD = 12.24$ ). Forty (18 females) participants were located in India with a mean age of 30.68 years ( $SD = 11.56$ ). Participants were selected at random.

### 2.2 Materials and Stimuli

In a questionnaire format, participants were given 50 algebra questions, followed by a short break of 2 minutes. For example, one question asked, "If  $x = 9$ , then  $6x + 1 = 55$ ". Participants were asked to determine if the questions were true or false. Half of the questions were correct, while the other half were incorrect. Following the break, participants were given the same 50 questions in another block of trials, for a total of 2 blocks of trials (required for a PPT analysis). In each block of trials, the questions were presented in a random order. Following this, participants were asked about their age, gender, what country they lived in, ethnicity, and how many math courses they had taken since high school. Participants were then debriefed and dismissed.

### 2.3 Design

All participants were given the same questions. Indian and American participants self-identified during the process, and were put into separate groups for analysis.

## 3. Results

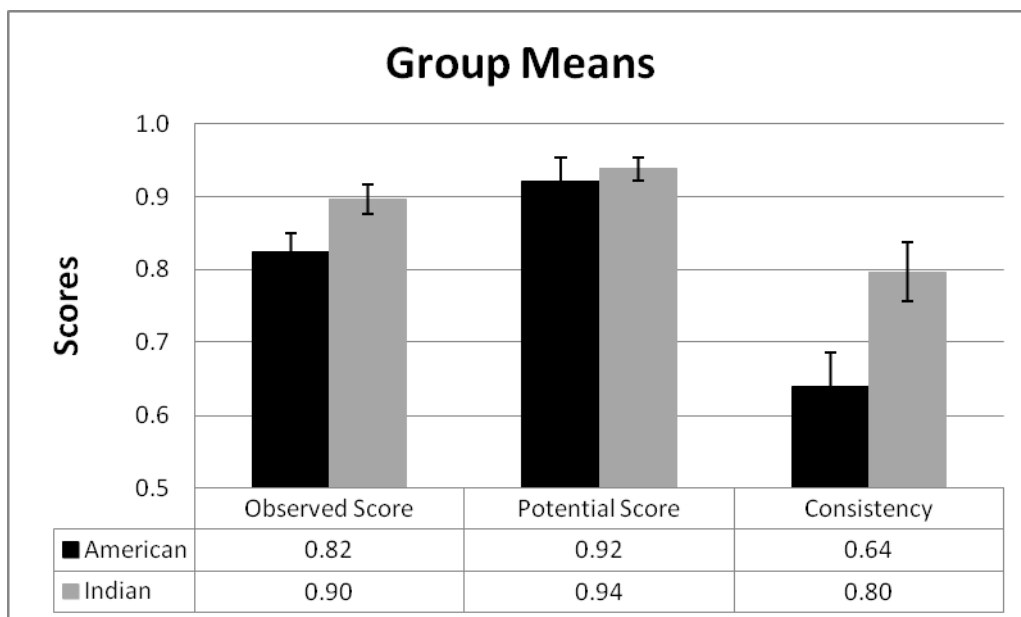


Figure 1. Data from the experiment. SE bars are included.

Using the PPT methodology described in the Introduction, the observed scores were converted into potential scores. Consistency values were also calculated during the process. These data are presented in Figure 1. A two-way ANOVA using Scores as a within-participants factor and Country as a between-participants factor found a significant main effect of Scores,  $F(1, 78) = 24.41, p < .001, \text{partial-eta squared} = .24$ , and a significant interaction between Scores and Country,  $F(1, 78) = 4.00, p < .05, \text{partial-eta squared} = .05$ . There was no main effect of Country,  $F(1, 78) = .08, p = .15, \text{partial-eta squared} = .03$ . Importantly, the interaction indicated that the difference in observed scores was greater than the difference in potential scores. Additionally, the difference between the consistency scores was significant,  $t(78) = 2.59, p = .01$ ; two-tailed.

On average, the American participants had taken 4.1 math courses since graduating from high school. This was significantly lower than the average amount for Indian participants (7.9 math courses),  $t(78) = 2.50, p = .01$ ; two-tailed.

#### 4. Discussion

A careful look at Figure 1 makes it obvious that Indians outperformed Americans. Was this because Indians had greater potential scores (more favorable systematic factors such as knowledge of how to work algebra problems) or they were simply more consistent? Although Figure 1 makes clear that Indians had slightly better potential scores, the main factor responsible for better Indian than American performances was that Indians were more consistent. Put another way, there was less randomness in the Indian than the American responses. This is a startling finding as many researchers have attributed differences in mathematics scores to systematic factors rather than to a difference in randomness. We also noted that the sample from India, on average, had more mathematics classes than did the sample from the US. This likely means that Indians perceived the items as easier than Americans did. Several PPT articles show that item difficulty primarily influences consistency more than potential performance (e.g. Hunt et al., in press; Rice et al., 2012; Rice, et al., 2011), and so the present findings and previous PPT findings inform each other.

There is literature demonstrating that adult Indians in university outperform Americans on tests of mathematical ability. This is not surprising given the great emphasis on mathematics training in India as students prepare for their MBA training. The present data are consistent with the literature as we also found that the Indian participants had significantly more mathematics training than did the American participants. But the present data extend the literature by testing whether the effect of the increased emphasis on mathematics training has its influence primarily through increasing the favorability of the systematic factors that influence algebra performance or through increasing performance consistency. Our findings demonstrate that the difference in mathematics emphasis in the two cultures affects algebra performance mainly through influencing performance consistency. Put in more general terms, the present findings narrow down the class mechanisms by which training in mathematics influences algebra performance because we now know that the effect is through consistency relevant variables rather than through systematic ones.

Of course, algebra is only a subset of mathematics, and so researchers should be cautious about generalizing the present findings to other areas in mathematics, and especially to other topics such as history, astronomy, and so on. In addition, a comparison between Americans and Indians is only one of a large set of comparisons that could be made, and so it is premature to assume that all cultural differences in observed performance are due more to differences in consistency than to differences in potential scores. Notwithstanding these cautionary notes, the present research opens up a new avenue for research in education. Specifically, having found that at least one cultural difference in observed performance, in one topic area, is due more to differences in consistency than to differences in potential performance, it suggests that more such effects can be found with additional cultures and topic areas. We hope and expect that researchers will explore this possibility with many cultures and topic areas.

#### 5. Conclusion

In conclusion, although previous literature has established the increased emphasis on mathematics training in India relative to the US, and that Indians outperform Americans on mathematics tests, the mechanism through which increased emphasis on mathematics training influences observed mathematics performance has yet to be explicated. The present findings address this issue. We showed that at least with respect to algebra, the effect is through consistency rather than through systematic variables. We anticipate that based on the present demonstration, future researchers will search for and discover exactly what the variables are that contribute to the decreased consistency in performance that we now know is associated with a lack of emphasis on mathematics training.

## References

- Hunt, G., Rice, S., Trafimow, D., & Sandry, J. (in press). Using Potential Performance Theory to analyze systematic and random factors in enumeration tasks. *American Journal of Psychology*.
- Phelps, R. (2001). Benchmarking to the World's Best in Mathematics: Quality Control in Curriculum and Instruction Among the Top Performers in the TIMSS. *Evaluation Review*, 25(4), 391-439. <http://dx.doi.org/10.1177/0193841X0102500401>
- Rice, S., & Trafimow, D. (in press). Time pressure heuristics can improve performance due to increased consistency: A PPT methodology. *Journal of General Psychology*.
- Rice, S., Geels, K., Hackett, H., Trafimow, D., McCarley, J. S., Schwark, J., & Hunt, G. (2012). The harder the task, the more inconsistent the performance: A PPT analysis on task difficulty. *Journal of General Psychology*, 139(1), 1-18. <http://dx.doi.org/10.1080/00221309.2011.619223>
- Rice, S., Geels, K., Trafimow, D., & Hackett, H. (2011). Our students suffer from both lack of knowledge and consistency: A PPT analysis of test-taking. *US-China Education Review*, 1(6), 845-855.
- Rice, S., Trafimow, D., & Hunt, G. (2010). Using PPT to analyze sub-optimal human-automation performance. *Journal of General Psychology*, 137(3), 310 - 329. <http://dx.doi.org/10.1080/00221301003645236>
- Rice, S., Trafimow, D., Keller, D., Hunt, G., & Geels, K. (2011). Using PPT to correct for inconsistency in a speeded task. *Journal of General Psychology*, 138(1), 12-34. <http://dx.doi.org/10.1080/00221309.2010.531791>
- Trafimow, D., & Rice, S. (2008). Potential Performance Theory: A general theory of task performance applied to morality. *Psychological Review*, 115(2), 447-462. <http://dx.doi.org/10.1037/0033-295X.115.2.447>
- Trafimow, D., & Rice, S. (2009). Potential Performance Theory (PPT): Describing a methodology for analyzing task performance. *Behavior Research Methods*, 41(2), 359-371. <http://dx.doi.org/10.3758/BRM.41.2.359>
- Trafimow, D., & Rice, S. (2011). Using a sharp instrument to parse apart strategy and consistency: An evaluation of PPT and its assumptions. *Journal of General Psychology*, 138(3), 169-184. <http://dx.doi.org/10.1080/00221309.2011.574173>
- Trafimow, D., Hunt, G., Rice, S., & Geels, K. (2011). Using potential performance theory to test five hypotheses about meta-attribution. *Journal of General Psychology*, 138(2), 1-13. <http://dx.doi.org/10.1080/00221309.2010.540591>
- Trafimow, D., MacDonald, J., & Rice, S. (in press). Using PPT to account for randomness in perception. *Attention, Perception and Psychophysics*.

## Appendix A (taken from Rice & Trafimow, in press)

The PPT strategy is as follows. In a dichotomous task, there are two possible responses to each trial, and each possible response might be correct or incorrect. Thus, for each participant, it is possible to create a 2 (option chosen) by 2 (option correct) frequency table. Table 1 illustrates the frequency table with the four cells. The participant can choose the first option and be correct (cell *a*) or incorrect (cell *c*), or the participant can choose the second option and be correct (cell *d*) or incorrect (cell *b*). There are also margin frequencies constructed as follows. The first row frequency is the sum of cells *a* and *b* whereas the second row frequency is the sum of cells *c* and *d* ( $r_1$  and  $r_2$ , respectively). The first column frequency is the sum of cells *a* and *c* whereas the second column frequency is the sum of cells *b* and *d* ( $c_1$  and  $c_2$ , respectively). Each participant's frequency table can be converted to a correlation coefficient via Equation 1.

$$r = \frac{|ad-bc|}{(a+b)(c+d)(a+c)(b+d)} = \frac{|ad-bc|}{r_1 r_2 c_1 c_2} \quad (1)$$

The second step, assuming that there were at least two blocks of trials in the experiment, is to obtain the consistency coefficient. This is simply the correlation, for each participant, across blocks of trials.

The third step is to use the famous correction formula that was originally derived by Spearman (1904), is a standard derivation from classical true score theory, but also can be derived from more modern theories (see Allen & Yen, 1979; Cohen & Swerdlik, 1999; Crocker & Algina, 1986; Gulliksen, 1987; Lord & Novick, 1968 for reviews). The formula is provided in Equation 2, where  $R$  is the best estimate of the correlation coefficient that would be obtained in the absence of randomness and  $r_{xx'}$  is the consistency coefficient across blocks of trials that was computed on Step 2.

$$R = \frac{r}{\sqrt{r_{xx}}} \quad (2)$$

It is worth noting that Equation 2 is a special case of a more general correction formula. In the general case, there are two variables being correlated, and so there is a consistency coefficient associated with each variable. However, we are interested in only one variable; this is participant's responding to the task trials. Therefore, for our present purposes, the simplified formula is sufficient. But it is possible to imagine a task where there is no correct answer, and so the criterion for "success" is agreement with another person. In this case, there would be a consistency coefficient for each person because each person would be capable of responding with some degree of randomness. As Trafimow and Rice (2008) showed, when there is no correct answer, the more general correction formula must be used, where each person's responding is treated as a separate variable.

Given that a corrected correlation coefficient has been obtained, the next step is to convert it back into a frequency table. But the new frequency table is not a table of observed frequencies, but rather of *potential* frequencies, which are the best estimates of the frequencies that would be obtained in the absence of randomness (i.e., with perfect consistency). These can be obtained via Equations 3-6 below. Equations 3-6 require fixing the margin frequencies at the obtained levels, similar to a Chi Square test or a Fisher's Exact test, and using  $R$  to estimate the potential cell frequencies. Trafimow and Rice (2008) provide derivations of Equations 3-6 and Trafimow and Rice (2009) provide empirical tests that strongly validate the assumptions. Finally, Trafimow and Rice (2011) provide further validation in the form of computer simulations. Consistent with previous research, we use upper case letters ( $A$ ,  $B$ ,  $C$ ,  $D$ ,  $R_1$ ,  $R_2$ ,  $C_1$ ,  $C_2$ ) to refer to the cell and margin frequencies in Equations 3-6 below.

$$A = \frac{R\sqrt{R_1R_2C_1C_2} + C_1R_1}{(R_1 + R_2)} \quad (3)$$

$$B = \frac{R_1(R_1 + R_2) - (R\sqrt{R_1R_2C_1C_2} + C_1R_1)}{(R_1 + R_2)} \quad (4)$$

$$C = \frac{C_1R_2 - R\sqrt{R_1R_2C_1C_2}}{(R_1 + R_2)} \quad (5)$$

$$D = \frac{C_2(R_1 + R_2) - [R_1(R_1 + R_2) - (R\sqrt{R_1R_2C_1C_2} + C_1R_1)]}{(R_1 + R_2)} \quad (6)$$

The foregoing steps for using PPT can be summarized as follows, providing that the participants have responded to two blocks of matched trials. First, the researcher uses Equation 1 to convert each participant's observed performance table into a correlation coefficient. Second, the researcher obtains a consistency coefficient for each participant. Third, the researcher uses Equation 2 to obtain a corrected or potential correlation coefficient. Finally, the researcher uses Equations 3-6 to estimate the cell frequencies that would be obtained without randomness (i.e., potential cell frequencies). Once the potential cell frequencies have been obtained, each participant's potential performance, or the performance level each person would have if the responses had been perfectly consistent across blocks can be calculated via Equation 7 below.

$$\text{potential performance} = \frac{A+D}{(A+B+C+D)} \quad (7)$$