

April 23, 2014

To:

Nathan Kuncel

Distinguished Professor of Psychology, University of Minnesota kunce001@umn.edu

Lee Alan Branum-Martin

Associate Professor, Department of Psychology and Co-Investigator at the Center for the Study of Adult Literacy at Georgia State University branummartin@gsu.edu

Dennis Kramer

Assistant Professor of Higher Education at the University of Virginia  
dkramer@virginia.edu

I'm afraid that you guys got snookered. You may be interested in some history which may be found at  
<http://paperclassinc.com/jay-smith-mary-willinghams-enemies/>

When I heard that our Provost was going to seek a review of Mary Willingham's work, I offered to do a review myself. I would have done an honest independent review as I am sure you tried to do. The difference is that I would have insisted on much more information than you were provided and I would have had a discussion with Willingham about her claims.

The Provost states in his "Executive Summary":

"Not all of the SATA subtests were used for the screenings. For example, the SATA Reading Comprehension and Writing Composition subtests were not used for screening".

This is NOT correct; this statement is true only of the INITIAL screening but not the followup. The questions he asked you to address were based on a misunderstanding of what Mary Willingham did. I attach a preliminary report which you were NOT given and which clarifies this - "The Incidence of Learning Disabilities and ADHD in Freshman Student Athletes". I am confident that the University was provided a copy of the report. The report says

"Forty-six (39%) student athletes were identified as having possible ADHD and/or a LD based on the results of the screening. Thirty-six (30%) were subsequently diagnosed as a result of a complete psycho-educational evaluation (Wechsler Adult Intelligence Scale-IV and completion of the SATA)." The SATA diagnostic test was administered by Lynda Johnson, a UNC-hired Ph.D. psychologist who was the coauthor of this report.

In the executive summary which you are said to have approved, the Provost states:

"The roughly 182 athletes in revenue sports (generally considered football, men's basketball and women's basketball) who were mentioned in the public claims represent

a small fraction of approximately 1,800 student-athletes who attended UNC in the eight-year period between 2004 and 2012,"

This is true but the concern is with these 182 primarily revenue sport athletes, not with college athletes in general. In this regard, the Faculty Council minutes state:

"Mr. Farmer said that when using CNN's threshold for 1,377 first year student-athletes admitted between 2004 and 2012, ... When limiting the data for student-athletes to football, men's basketball and women's basketball, 90 percent of 341 first-year student-athletes met CNN's threshold. Of the 34 students who did not, 20 have graduated or remain enrolled in good standing, 10 left academically ineligible to return, <http://tinyurl.com/l38vgb8>

There are two facts that we know for certain. -

36 of 117 athletes initially reviewed were diagnosed as ADHD; the University has not released the figures for the full 182 athletes cited by Willingham.

34 of 341 athletes in revenue sports had SAT reading scores (or equivalent ACT scores) less than 400 where 200 is the minimum score. This is far below the UNC 25th percentile of 590. In light of what we now know about "paper classes", plagiarism, excessive help in writing papers, and strenuous efforts to keep athletes eligible, the fact that "20 have graduated or remain enrolled in good standing" must be taken with a grain of salt. This does not reflect well on the University's past admission policies for athletes in revenue sports.

There are obvious flaws in Willingham's research; her sample of 182 athletes is not a random sample of the 341 athletes reported by Farmer. We do not know how she arrived at her grade-level statements though we do know that they were not based solely on a vocabulary test but were based on the full SATA, the SAT, the WAIS and her personal judgment. The 60% figure cited by CNN may be due to a misinterpretation of her analyses. The preliminary report cited above suggests that the correct figure may be closer to 39%, the proportion diagnosed as LD/ADHD. It is certainly true that Willingham should have and could have obtained IRB approval to collect information using athletes' names. She need only have provided the justification and agreed to keep the names confidential as she has done.

#### Comment on the "Executive Summary"

Page 4 refers to an "internal review team" without naming the members or indicating their competence. The comments on this page are without merit since the Provost did not understand that Willingham's public statements were not based on the initial screening but on the followup which included the full SATA, SAT, and the WAIS intelligence scale. As an example of their misunderstanding, the faculty council

minutes state that:

"Provost Dean explained that the claims about reading levels were based on confusion between the standard score and grade equivalent."

The suggestion that Mary Willingham does not understand the difference between standard scores and grade equivalents is simply absurd, given her background.  
<http://www.unc.edu/~willingh/2013resume.pdf>

Willingham has stated and written that her judgments were based on the full SATA, SAT, and the WAIS intelligence scale as well as her personal observations.

The questions posed to you by the Provost are certainly unartfully phrased and could be interpreted as revealing the bias of the Provost and his associates. As a result, you reviewers chose to ask your own questions. For example, Branum-Martin writes

"The framing of this question is dangerously simplistic. ... A more reasonable question would be ...

Since none of you have answered the specific questions posed by the Provost, I will answer them for you:

1. Is the SATA Reading Vocabulary (RV) subtest a true reading test?

Of course a Vocabulary test is not a reading test and Willingham never claimed that it was; As you well know, it certainly is related to reading ability.

2. Were SATA RV standard scores mistakenly interpreted as grade equivalents to report on UNC student-athletes' reading grade levels?

No; this is nonsense.

3. Can you verify that the SATA Reading Vocabulary (RV) data in the January 13 data set is in the form of standard scores?

Any undergraduate psychology student could.

4. Can you verify that SATA RV standard scores are not the same as grade equivalents?

Any undergraduate psychology student who has taken a course in test theory could.

5. The SATA test is normed against a sample of 1005 people with the following demographics: ... Based on your professional expertise and your analysis of the data set, would this difference in demographics have an impact on reporting results?

I agree with Professor Kuncel who wrote:

"the SATA norms appear to be adequate although based on somewhat small samples by age group (a comment echoed in Raju's critique of the SATA). The sampling plan was somewhat driven by convenience but a good effort was made to get a representative sample. It reasonably reflects the general population with respect to gender, geographic region, race and ethnicity. ...

If the goal is to examine how the tested student athletes compare with the general population, then the SATA norm group is perfectly reasonable"

This is certainly the goal; it is NOT to compare UNC athletes with those at other schools.

6. To what extent should the SATA RV subtest be considered a measure of literacy? Should these results be reported in terms of grade level?

I agree with Professor Kuncel who wrote:

"The Reading Vocabulary (RV) scales assess a person's vocabulary knowledge. It cannot be viewed as a reasonable or comprehensive assessments of adult literacy. ... if a person does well on one assessment (e.g., vocabulary tests) they are likely to do well on another assessment (e.g., college admissions tests). Similarly, if a person struggles with one they are likely to struggle with another. ... This relationship is consistent with the use of the Reading Vocabulary measure as a screening tool to identify students who may need additional academic support or have a learning disability."

Of course, Willingham relied on the full SATA as well as the SAT and WAIS.

#### Other comments by the outside experts

"the majority (n=109) of the 176 students with valid scores in the sample had Reading Vocabulary grade equivalents above 12th grade."

This means that 38% had grade equivalents below the 12th grade. Of course, this statement was based solely on the Vocabulary test, but it is hardly a ringing endorsement of UNC admission policies.

"Standard scores for the SATA, which range from 1 to 20, do not represent grade level any more than SAT scores."

Of course not and Willingham did not make such an absurd mistake; this is the Provost's public claim.

"Reading ability should not be reported as grade equivalents."

"Grade equivalents should never be presented as a measure of performance."

These are overstatements; grade equivalents (which were not based solely on the SATA vocabulary test) are more appropriate for a non-technical audience. For example the College Board gives the average score for 11th graders on the PSAT to help students understand what their PSAT score means.

<http://www.collegeboard.com/student/testing/psat/scores.html>

Although you chose not to answer the questions posed by the Provost, you did ask some interesting questions and do some interesting analyses yourselves. For example Professor Branum-Martin poses the question:

Specific Question 3: Based on your professional expertise, is it possible to assign a reading grade level to a student based on a combination of SATA RV and writing subtests, SAT verbal scores, and one on one work with students? The answer is:

"Yes, such an assignment would be possible, ...Having other data such as SAT verbal scores could provide helpful information in estimating students' practical levels of performance. However, such scores were not provided in the given spreadsheet. One-on-one work with students involves professional opinion, but could hopefully be corroborated with a variety of objectively scored tests, performance assessments, and supportive documentation from professionals familiar with the students (e.g., teachers, tutors)." She also says "As noted previously, vocabulary is essential, but not sufficient for reading comprehension. In this respect, the SATA Reading Vocabulary subtest could be informative regarding reading related skills, especially word knowledge, but may not be sufficient for broad judgments of student literacy. Ideally, other sources of information should be included, especially tests of reading comprehension."

Of course Mary Willingham DID have such information as does the University. Furthermore, as a tutor, Willingham had extensive familiarity with many of the students. Branum-Martin further says "Reading comprehension scores were not provided. The full records of the SATA items and scores and other reading measures should be obtained for the sake of serving the student athletes and for the evaluation of the programs which serve those athletes." Both Willingham and the University had such information but it was not provided to any of you.

Professor Kramer's analyses are interesting but are largely irrelevant since they are based solely on the Vocabulary scale while Willingham's conclusions were based on those who took the entire SATA as well as the SAT and WAIS. Of course, this is not his fault; he worked with what he was given. He also expresses concern about "the overrepresentation of revenue-generating student-athletes within the sample" but these are the athletes of primary concern.

In conclusion, I have one question for you; do you really think that it should ever have been acceptable for a university with the stature of UNC-CH to admit students with SAT scores below 400 as was done in the case of 34 athletes in the years under study?

What would you regard as a minimally acceptable score? I would be surprised if any non-athlete had a score as low as 450. I am strongly in favor of our university having strong athletic programs but not at the expense of accepting student who are unlikely to succeed or who can graduate only with their being guided into phony courses and receiving unacceptable help with their studies.