Student Academic Motivation and Non-Cognitive Skills: Improving Comparability across Cultures and Gender with the Anchoring Vignettes Method

Hanka Vonkova, Ondrej Papajoanu, Katerina Kralova
Charles Univesity, Prague

Gema Zamarro
University of Arkansas

Word Count: 1,000

Background

Student academic motivation is an important determinant of school performance (e.g. Fortier et al., 1995) and one of the key concerns of both teachers and parents worldwide. Most of the research, however, lacks an international perspective and it has ignored the fact that students in different cultures may have a different understanding of concepts such as effort and academic motivation (Elliot and Bempechat, 2002).

One important challenge of international and group comparisons of the academic motivation of students is that comparisons of self-reported measures can be biased if respondents differ in their use and interpretation of the different scales in the provided self-reported questions. This issue, which has been referred in the literature as reference group bias, is an important problem that has been mostly ignored in most education research. Building on our prior work, described in Vonkova et al. (2018), we study the use of the anchoring vignettes method for enhancing comparability of student self-reports of academic motivation across countries and across groups defined by gender.

Related recent literature emphasizes the role that non-cognitive or character skills related to conscientiousness (willingness to work hard), self-control, and perseverance have in determining students' academic success and STEM outcomes (see, Almlund et al., 2011; Weel, 2008; Heckman, Stixrud, and Urzua, 2006; Aryee, 2017). Importantly, these non-cognitive skills have been shown to be malleable (see e.g. Alan, Boneva, and Ertac, 2019) and they could mediate the effect of academic motivation on student outcomes (see, e.g. Richardson and Abraham, 2009). However, measuring non-cognitive skills can be also challenging due to the fact that self-reports of these skills can also be affected by multiple types of bias (e.g. reference group bias and social desirability bias; see Duckworth and Yeager, 2015). An alternative approach recently proposed in the literature is to use measures of survey and test effort as alternative measures of non-cognitive skills related to conscientiousness (Soland et al., 2019).

In this project, we also study how reported achievement motivation relates to task-based measures of students' non-cognitive skills related to conscientiousness based on the effort put forward on the survey and test that are part of the PISA assessment, and whether these non-cognitive skills could mediate the observed relationships between academic motivation and student STEM outcomes.

Research Questions

1) Can anchoring vignettes help improve comparability of self-reported measures of academic motivation across countries and within countries by gender?
2) How does academic motivation relate to Math and Science interest and performance? Does this relationship vary across countries? Does it vary by gender within countries?
3) How does academic motivation relate to performance-task measures of student's non-cognitive skills related to conscientiousness based on the effort put forward in the PISA assessment? Do these non-cognitive skills mediate the relationship between academic motivation and STEM outcomes?

Data Collection and Analysis

Our project uses data from PISA (2015). PISA is a triannual survey which evaluates how well 15-year-old students are capable of using their knowledge and skills to meet real-lifechallenges in the areas of mathematics, reading, and science. The number of participants in 2015 was about 570,000 students from 72 countries and economies.

PISA (2015) is ideal for our project. Firstly, the student questionnaire included measures of student academic motivation along with anchoring vignettes. To correct self-reported measures of students' motivation, we follow Vonkova et al. (2018) and employ the parametric model of the anchoring vignettes method, introduced by King et al. (2004) as the Compound Hierarchical Ordered Probit (CHOPIT) model. Secondly, PISA (2015) was the first time that students were asked to perform the assessment in the computer and a result, we have information on response times for each question in the PISA test. Item response times are used to identify rapid guessing rates, i.e. instances of responses that were provided so quickly that the item's content probably could not have been understood (Lee & Jia, 2014; Rios et al., 2016; Wise & Kong, 2005). Finally, the PISA test is followed by the PISA student questionnaire which we use to build measures of survey effort. In particular, we follow Hitt et al. (2016) and Zamarro et al. (2018) to build item non-response rates, i.e. proportion of questions that students skipped out of the total of questions that they were asked to answer.

For our STEM-related outcomes, we use information about math and science performance in the PISA assessment and self-reported student interest in science. We then study the relationship between academic motivation and STEM outcomes following this model:

$$STEM\_Outcome_i = \beta_0 + \beta_1 X_i + \beta_2 female_i + \beta_3 Academic\_Motivation_i + \beta_4 Academic\_Motivation * female_i + \varepsilon_i$$

We run separate analysis for each country or economy included in the PISA 2015 assessment. $Academic\_Motivation_i$ is our main variable of interest. We perform separate regressions including raw self-reports of this variable, as well as corrected reports using the anchoring vignettes method. This allow us to assess the importance for our results of correcting for the differential use of reporting scales across students. We pay special attention to understanding the source of gender differences in STEM outcomes by including a gender dummy variable in the analysis as well as interactions between student's academic motivation and gender. Finally, X_i include important socio-economic controls like immigrant status, socio-economic status of the family, type of school the student is attending (i.e. public, private), urbanicity, or whether the student has been retained. Standard errors are clustered at the school level to take into account the fact that we have multiple students in the sample attending the same schools.

To study the mediating role of students' conscientiousness, proxied through survey and effort, we add an interaction term between academic motivation and different measures of students' effort in the PISA assessment.

Results

Our results so far highlight the importance and potential of anchoring vignettes methods to improve the comparability of student motivation measures across cultures and gender groups.

Conclusions

This project's contribution to the literature is twofold. Firstly, we provide a methodological contribution improving the comparability of academic motivation measures by using anchoring vignettes methods. Secondly, we provide an empirical contribution by studying the relationship between academic motivation, non-cognitive skills related to conscientiousness, and science-related outcomes.

References

Alan, S., Boneva, T., & Ertac, S. (2019). Ever failed, try again, succeed better: Results from a randomized educational intervention on grit. *The Quarterly Journal of Economics*, 134(3), 1121-1162.

Almlund, M., Duckworth, A. L., Heckman, J. J., & Kautz, T. D. (2011). Personality psychology and economics. *Handbook of the Economics of Education*, 4, 1-18.

Aryee, M. (2017). College students' persistence and degree completion in science, technology, engineering, and mathematics (STEM): The role of non-cognitive attributes of self-efficacy, outcome expectations, and interest. Ph.D., Seton Hall University.

Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher*, 44(4), 237–251.

Elliott, J. G., & Bempechat, J. (2002). The culture and contexts of achievement motivation. *New Directions for Child and Adolescent Development*, 2002, 7–26.

Fortier, M. S., Vallerand, R. J., & Guay, F. (1995). Academic motivation and school performance: Toward a structural model. *Contemporary Educational Psychology*, 20, 257-274.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3), 411-482.

Hitt, C., Trivitt, J., & Cheng, A. (2016). When you say nothing at all: The predictive power of student effort on surveys. *Economics of Education Review*, 52, 105–119.

King, G., Murray, C., Salomon, J., & Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review*, 98, 567–583.

Lee, Y. H., & Jia, Y. (2014). Using response time to investigate students' test-taking behaviors in a NAEP computer-based study. *Large-Scale Assessments in Education*, 2(1), 8.

Richardson, M., & Abraham, C. (2009). Conscientiousness and achievement motivation predict performance. *European Journal of Personality*, 23(7), 589-605.

Rios, J. A., Guo, H., Mao, L., & Liu, O. L. (2016). Evaluating the impact of careless responding on aggregated-scores: To filter unmotivated examinees or not? *International Journal of Testing*, 1–31.

Soland, J., Zamarro, G., Cheng, A., & Hitt, C. (2019). Identifying Naturally Occurring Direct Assessments of Social-Emotional Competencies: The Promise and Limitations of Survey and

Assessment Disengagement Metadata. *Educational Researcher*. https://doi.org/10.3102/0013189X19861356

Vonkova, H., Zamarro, G., Hitt, C. (2018). Cross-country heterogeneity in students' reporting behavior: The use of the anchoring vignette method. *Journal of Educational Measurement*, 55(1), 3-31.

Weel, B. T. (2008). The noncognitive determinants of labor market and behavioral outcomes: Introduction to the symposium. *Journal of Human Resources*, 4 3(4), 729–737.

Wise, S. L., & Kong, X. (2005). Response time effort: A new measure of examinee motivation in computer-based tests. *Applied Measurement in Education*, 18(2), 163–183.

Zamarro, G., Cheng, A., Shakeel, M. D., & Hitt, C. (2018). Comparing and validating measures of non-cognitive traits: Performance task measures and self-reports from a nationally representative internet panel, *Journal of Behavioral and Experimental Economics*, 72, 51-60.