

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/278038022>

Determination variables which define Turkish students' PISA 2009 reading literacy with chaid analysis

Article in *Pakistan Journal of Statistics* · December 2014

CITATION

1

READS

168

2 authors:



[Asiye Şengül Avşar](#)

Recep Tayyip Erdoğan Üniversitesi

18 PUBLICATIONS 27 CITATIONS

[SEE PROFILE](#)



[Ezel Tavşancıl](#)

Ankara University

25 PUBLICATIONS 218 CITATIONS

[SEE PROFILE](#)

**DETERMINATION VARIABLES WHICH DEFINE TURKISH STUDENTS'
PISA 2009 READING LITERACY WITH CHAID ANALYSIS***

Asiye ŞENGÜL AVŞAR¹ and Ezel TAVŞANCIL²

¹ Department of Educational Measurement and Evaluation,
Recep Tayyip Erdoğan University, Turkey.

Email: asiye.sengul@erdogan.edu.tr

² Department of Educational Measurement and Evaluation,
Ankara University, Turkey. Email: etavsancil@ankara.edu.tr

ABSTRACT

The aim of this study to determine the variables explaining reading literacy of Turkish students through their responds to PISA 2009 questionnaire. Variables explaining Turkish students' reading literacy were determined with items which were selected from PISA student questionnaire. The data were analyzed by using CHAID (Chi-squared Automatic Interaction Detection) analysis which is one of the algorithms of data mining decision trees. The result of the study indicated that the best defining variable of Turkish students' reading literacy is to possessions literature.

KEYWORDS

PISA 2009; Turkish Students' Reading Literacy; CHAID Analysis.

1. INTRODUCTION

Turkey has joined the studies which are done internationally to compare the current education system with the other countries' and to evaluate the students' achievement in international dimension beside of evaluating the students' achievement in national dimension. One of the studies that Turkey joined in international dimension is Programme for International Student Assessment (PISA). PISA 2009 of which the focus of reading literacy is important in terms of Turkey's realizing the 5-year-development of primary school program which was renewed by affecting the results of PISA 2003 and evaluating the products of renewed primary school program.

Reading comprehension skill is a virtual life skill for individuals. Unless this skill is developed, individuals loading a meaning to the events occurred in the society, associating them, judging by thinking on these events and creating a solution will not be possible (Kutlu, 2004). With the development of reading literacy which is the basic of all lessons in the primary school, students will become individuals whose thinking sense and comment power develops, who gain feeling and admiration delicacy, starts recognizing

* This study is the summary of a master thesis submitted to Ankara University Department of Measurement and Evaluation; and presented in 1st International Eurasian Educational Research Congress in Istanbul, Turkey, 24-26 April 2014

national and global cultural accumulation (Aslanoğlu, 2007). So, according to the PISA 2009 results, it is considered that determining the variables which describes the Turkish students' reading literacy provides both the development of other skills of the students' as well as giving an opportunity for enhancing the reading literacy. Searching the variables that describe the reading literacy of Turkish students in PISA 2009 project is considered as important in the terms of taking necessary precautions to develop not only reading literacy but also mathematics and science literacy. Because of the limited studies in which Turkish students' reading literacy are searched with PISA data, this study is considered to provide benefits to literature.

For these reasons it is found necessary that Turkish students' reading literacy in PISA 2009 is determined with CHAID analysis which is one of the decision trees algorithms of reading literacy. In this research in which determining the variables that describe the reading literacy of Turkish students was aimed, answers were looked for the questions: "What is the independent variable which divides current data set into homogeny sub groups and describes best the reading literacy of Turkish students?" and "What are the other variables that describe the reading literacy of Turkish students?"

2. RESEARCH DESIGN AND METHOD

The search is in the survey design. PISA 2009 was implemented in a stratified sampling method at 56 cities from 12 statistical region units and 4996 students from 170 schools. Hard effort and sources are spent for validity and reliability of instruments which are used in PISA in all countries and for minimizing the cultural and lingual differences. There are some safety assurance which is mandatory to obey for the process of translation, sample and gathering data (MEB, 2010). In the students' questionnaire; there are some items related to themselves, their families, houses, reading activities, time that they spend to learn, class and school region, Turkish classes, libraries, strategies of reading and understanding a text.

In PISA application, the students response to the determined sub-group of the questions. Non-observed responses are predicted by the help of observed responses in PISA. Theoretical achievement distribution is created for each student. Five possible values which are chosen randomly from this distribution are recommended to be used in statistical process (EARGED, 2010). In this research, the arithmetic mean of five possible values that were calculated in reading literacy was calculated.

The variables that describe the reading literacy of Turkish students were determined with CHAID analysis which is one of the data mining method. CHAID analysis developed by Kass (1980) has doing the most suitable division in data set by using chi-square statistic (SPSS, 1998). CHAID analysis is an explanatory method which is used in studies to determine the relations between a dependent variable and a series independent variable (Doğan and Özdamar, 2003). This analysis is an algorithm which divides the current data set into detailed and private sections (Diepen and Franses, 2005).

CHAID analysis without being affected from missing values in data set, as it can divide the whole population into stable node with its strong iteration algorithm besides it gathers the missing values in separate group, a regression equation which will

be gained with this analysis is individualized from well-known classical assumptions (normality, linearity, homogeneity etc.) (Horner, Fireman and Wang, 2010; Kayri and Boysan, 2007).

3. RESULT

As the result of CHAID analysis the variable that describe the reading literacy of Turkish students best, that will divide current data set into homogeneity sub groups is seen in Figure 1. Also the variables that describe students reading literacy, homogeneity sub groups formed according to these variables, mean achievement scores of the students in these homogeneity sub groups, the number of students in these homogeneity sub groups, the percentage of the number of these students in all data set are given in Figure 1, 2 and 3.

In Figure 1, it is seen that the variable that describe the reading literacy of Turkish students is seen as possessions literature variable ($F_{(1, 4994)}=780.4907$; $p<0.05$). The mean achievement score of reading literacy of Turkish students is 465.7125, the proficiency level of reading literacy is second proficiency level. The ones who say 'no' to the question about possessions literature and give no answers (missing variables) to this question are gathered in "Node 1" and the ones who say 'yes' to the question above is gathered in "Node 2".

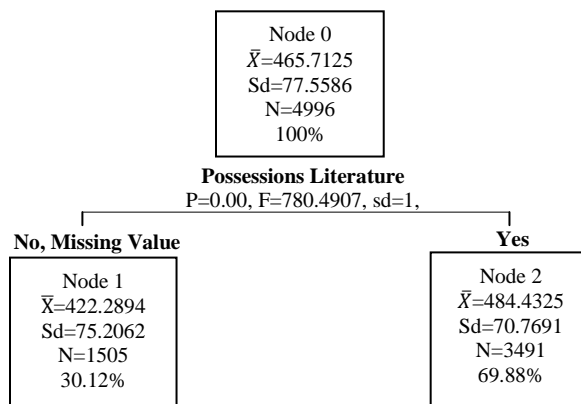


Figure 1: The Variable That Describe the Reading Literacy of Turkish Students' Best

In Figure 2, the variable that describe reading literacy of whom say 'no' and give no answers to the question about possessions literature is seen as grade level ($F_{(2, 1502)} = 246,8533$; $p < 0.05$). It is seen that, in Node 3 primary school students, in Node 4 9th grade students, in Node 5 10th, 11th and 12th grade students are gathered. The variable that describe reading literacy of primary school student in Node 3 has an item like 'When I study and I don't understand something, I look for additional information to clarify this' which is about studying from reading activities ($F_{(1, 109)} = 15.7410$; $p < 0.05$).

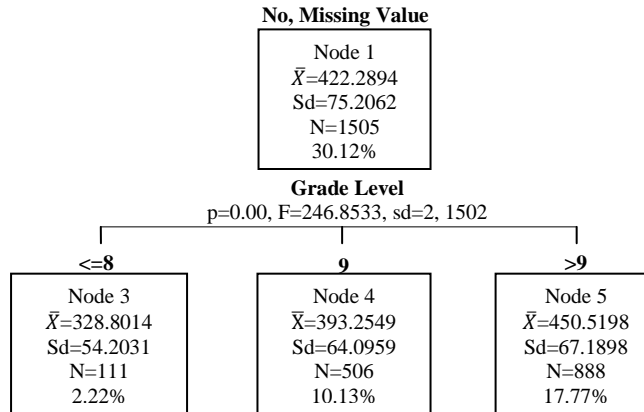


Figure 2: Other Variables that Describe Reading Literacy of Turkish Students' and Homogeny Sub-Groups that These Variables Divide

It was seen that the variable that describes reading literacy for 9th grade students is an item like 'I underline the important parts of the text' which is potential strategy to be used in case of understanding and remembering the text from the variables that are about the strategies of reading and understanding the text ($F_{(2, 503)}=24.9027$; $p<0.05$) and for 10th, 11th and 12th grade students is the school type ($F_{(3, 884)}=109.6024$; $p<0.05$).

In Figure 3, It is seen that the variable that describes the reading literacy of the students that answer as 'yes' to the question of possessions literature variable is the school type ($F_{(4, 3486)} = 482.2878$; $p<0.05$).

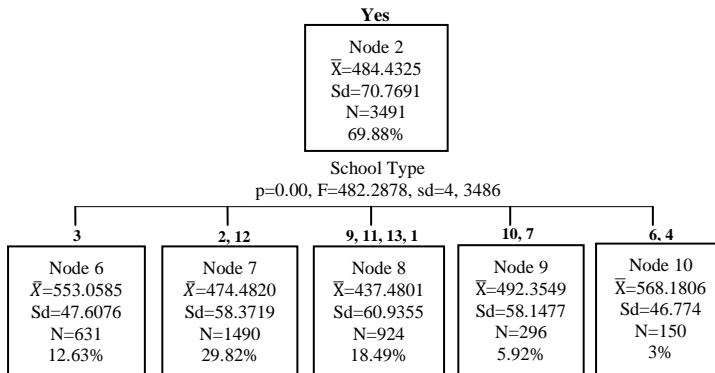


Figure 3: The Other Variables Describes the Reading Literacy of the Turkish Students' and Homogeny Sub-Groups that These Variables Divide

[Anatolian High School (3), General High School (2), Anatolian Technical High School (12), Vocational High School (9), Technical High School (11), Multi Programme High School (13), Primary School (1), Anatolian High School (10), Anatolian Fine Arts High School (7), Anatolian Teacher Training High School (6), Science High School (4)]

It was seen that, the variable which describes the reading literacy students who possession literature; for Node 6 is the gender variable ($F_{(1, 629)}=59.5622$; $p<0.05$); for Node 7 is the grade variable ($F_{(1, 1488)}=278.3409$; $p<0.05$); for Node 8 is gender variable ($F_{(1, 922)}=124.6942$; $p<0.05$); for Node 9 students is the time variable that the students spend time for studying to Mathematics, Turkish, Science and other lessons out of their school time ($F_{(1, 294)}=37.1536$; $p<0.05$) and for Node 10 students is gender variable ($F_{(1, 148)}=23.0390$; $p<0.05$).

4. CONCLUSION AND SUGGESTION

The variable that describes best the reading literacy of the Turkish students' is "possessions literature" variable. It was found that the reading literacy mean achievement score of the ones who possess literature are higher than the ones that do not. Also, the ones that answered as yes to this item took upper proficiency level than the ones that answered as no or left as unanswered according to the reading literacy proficiency level (third level) described in PISA. Students' reading literature is specified as a factor that develops their reading literacy. There are some researches in literature that the relationship between the reading habits of the students and academic achievement in different lessons. In these researches it is described that there is a high relation between the power of understanding what you have read and academic achievement (Güzeller, 2006; Keşan, Kaya and Yetişir, 2008; Kutlu, Yıldırım, Bilican and Kumandaş, 2010; Soylu and Tatar, 2006). When all the lessons at school should be considered to be studied, it cannot be thought that a student will be successful at lessons when he/she does not read carefully, does not understand what he/she reads. So, having and reading literature which is determined as an important factor on improving the reading literacy of the students should be supported.

According to the CHAID analysis results, it was seen that the reading literacy increased parallel to the increasing grade level. It is described in literature that although they are at different lesson field the upper grade students are more successful than the lower grade students (Güzel İş, 2009; Kotte, Lietz and Lopez, 2005). As regards to the CHAID analysis results, another variable which describes literacy is school type. School type separation has a very important place in Turkish Education system. The separation of the school type is very effective on the success of the students. The difference finding according to the school type that was gained as a result of research shows consistency by the researches done by Berberoğlu and Kalender (2005), Eraslan (2009), Yalçın and Tavşancıl (2014), Yılmaz (2009).

The most successful schools are determined as Anatolian Teacher Training High School, Science High School and Anatolian High School (fourth proficiency level) in the school type separation of the students who possess literature. There is serious difference between the school types in Turkey. As Berberoğlu and Kalender (2005) stated the school type differences are very remarkable as PISA results. In fact, the successful schools are one standard deviation of the international mean, unsuccessful one level down, there are two standard deviation between them. Therefore the difference between the school types should be decreased and opportunity equalization in education should be provided.

Another variable that describes the reading literacy of the students is gender. This finding of the research shows parallelization with the research findings which describes the students' success differences according to the gender (Kutlu et al., 2010; Sallabaş, 2008; Taube, Malin and Linnakyla, 2004; Thorpe, 2006).

As a summary, regarding to the result of CHAID analysis done, the primary students who do not possess literature and of them who sometimes searches for additional information and almost never searches for additional information for the subjects that they do not understand while studying are found the most unsuccessful students, the female students who studies in Anatolian Teacher Training High School and Science High School are found as the most successful students.

REFERENCES

1. Aslanoğlu, E.A. (2007). PIRLS 2001 Türkiye verilerine göre 4. sınıf öğrencilerinin okuduğunu anlama becerileriyle ilgili faktörler. Unpublished doctoral thesis. Ankara Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
2. Berberoğlu, G. ve Kalender, İ. (2005). Öğrenci başarısının yıllara, okul türlerine, bölgelere göre incelenmesi: ÖSS ve PISA analizi. *Eğitim Bilimleri ve Uygulama*, 4(7), 21-35.
3. Diepen, V.M. and Franses, H.F. (2006). Evaluating chi-squared automatic interaction detection. *Information Systems*, 31, 814-831.
4. Doğan, V. ve Özdamar K. (2003). CHAID analizi ve aile planlaması ile ilgili bir uygulama. *Tip Bilimleri Dergisi*, 23, 392-397.
5. EARGED (2010). Web: <http://earged.meb.gov.tr/pdf/pisa2009rapor.pdf>
6. Eraslan, A. (2009). Finlandiya'nın PISA'daki başarısının nedenleri: Türkiye için alınacak dersler. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi*, 3(2), 238-248.
7. Güzel, İ.Ş. Ç. (2009, Mayıs). Uluslararası öğrenci değerlendirme programında (PISA 2003) insan ve fiziksel kaynakların öğrencilerin matematik okuryazarlığına etkisinin kültürler arası karşılaştırılması. Paper presented at the 8. Ulusal Sınıf Öğretmenliği Sempozyumu, Eskişehir.
8. Güzeller, C. (2006). Ortaöğretim kurumları öğrenci seçme sınavının Türkçe dil yeterlikleri açısından modellenmesi. *Kastamonu Eğitim Dergisi*, 14(2), 403-412.
9. Horner, B.S., Fireman, D.G. and Wang, W.E. (2010). The relation of student behavior, peer status, race and gender to decisions about school discipline using CHAID decision trees and regression modeling. *Journal of School Psychology*, 48, 135-161.
10. Kayri, M. ve Boysan, M. (2007). Araştırmalarda CHAID analizinin kullanımı ve baş etme stratejileri ile ilgili bir uygulama. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 40(2), 133-149.
11. Keşan, C., Kaya, D. ve Yetişir, Ş. (2008). Türkçe-Matematik birlikteliğinin öğrenci başarısını etkileme gücü üzerine bir araştırma. *Üniversite ve Toplum Dergisi*, 8(2).
12. Kutlu, Ö. (2004). Türkiye'de Demokrasi anlayışının gelişmesini sağlayacak bir yol: okuduğunu anlama becerilerinin geliştirilmesi. Paper presented at the Uluslararası Demokrasi Eğitimi Sempozyum.

13. Kutlu, Ö., Yıldırım, Ö., Bilican, S. ve Kumandaş, H. (2011). İlköğretim 5. sınıf öğrencilerinin okuduğunu anlamada başarılı olup-olmama durumlarının kestirilmesinde etkili olan değişkenlerin incelenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 2(1), 132-139.
14. Kotte, D., Lietz, P. and Lopez, M.M. (2005). Factors influencing reading achievement in Germany and Spain: Evidence from PISA 2000. *International Education Journal*, 6(1), 113-124.
15. MEB (2010). Web: <http://www.earged.meb.gov.tr/pdf/pisa2009rapor.pdf>
16. Sallabaş, M. (2008). İlköğretim 8. sınıf öğrencilerinin okumaya yönelik tutumları ve okuduğunu anlama becerileri arasındaki ilişki. *İnönü Üniversitesi, Eğitim Fakültesi Dergisi*, 9, 16, 141-155.
17. Soylu, Y. ve Tatar, E. (2006). Okuma-Anlamadaki başarının matematik başarısına belirlenmesi üzerine bir çalışma. *Kastamonu Eğitim Dergisi*, 14(2), 503-508.
18. SPSS (1998). Answer Tree 2.0 User's Guide.
19. Taube, K. Malin, A. and Linnakyla P. (2004). Factors behind low reading literacy achievement. *Scandinavian Journal of Educational Research*, 48, 3.
20. Thorpe, G. (2006). Multilevel analysis of PISA 2000 reading results for the United Kingdom using pupil scale variables. *School Effectiveness and School Improvement*, 17(1), 33-62.
21. Yalçın, S. and Tavşancıl, E. (2014). The comparison of Turkish students' PISA achievement levels by year via data envelopment analysis. *Kuram ve Uygulamada Eğitim Bilimleri*, 14(3), 961-968.
22. Yılmaz, B.H. (2009). Turkish Students' Scientific Literacy Scores: A Multilevel Analysis of Data from Program for International Student Assessment. Unpublished doctoral dissertation, The Ohio State University, Columbus, Ohio.