

PS2CLH: A Learning Factor Model for enhancing students' ability to control their achievement

Arlindo Almada, London Metropolitan University, United Kingdom

Qicheng Yu, London Metropolitan University, United Kingdom

Preeti Patel, London Metropolitan University, United Kingdom

The Asian Conference on Education “Independence & Interdependence” (ACE2019)

Abstract

Numerous factors that influence students' academic performance involve issues beyond the individuals' control, such as national policies, government initiatives and university resources among many others. Even if students are aware of these factors, addressing them may be unfeasible. Identifying causes within students' control could both improve students' understanding of these factors as well as enabling students to independently deal with related issues. This paper proposes a student-controllable learning factor model that combines the perspectives of Psychology, Self-responsibility, Sociology, Communication, Learning and Health & wellbeing (PS2CLH). The proposed model used qualitative methods to identify underlying aspects affecting academic achievement and selected controllable factors.

This research reports on the outcomes of the employment of the PS2CLH model to predict student performance. Initially, data is collected through a self-evaluative web-based questionnaire. Each student's past performance and factors affecting this are then quantified. This study reveals the impact of students' controllable factors on student achievement. The model test results indicated 94% accuracy of successful prediction of the student performance based on the proposed PS2CLH model. The importance of “establishing and achieving personal goals” was higher than “stress”, “learning room” and “grammar and vocabulary” among other factors. This research raised participant students' awareness of PS2CLH perspectives, which helped them manage factors affecting academic performance more effectively. Consequently, most of the students have enhanced their academic performance by addressing these critical factors. However, due to the limitations of the current sample data, the PS2CLH model will be further monitored for various applications.

Keywords: Students performance, learning factors, grade prediction

Introduction

Studying the factors influencing a student's academic performance has been a matter of research for decades. Consequently, there is an overwhelming number of studies in this field. Among them, four researchers have presented the most influential work. Firstly, Professor John Hattie, one of the world's leading education researchers in this field. His ongoing research work *Visible Learning* focuses on the evaluation of learning and teaching techniques, models of measurement and performance indicators (Hattie, 2009). In 2018, Hattie's *Visible Learning* research synthesised findings from 1500 meta-analyses of 90,000 studies (Hattie, 2018). Secondly, Rossi and Montgomery's model focuses mainly on societal student's context, which points to two distinct scenarios. Firstly, the community environment and home quality, secondly the quality of the school such as the classroom conditions, curriculum and student's incentives (Akama, 2017). Thirdly, a research group led by Dunlosky from Kent State University in 2013 presented ten years of literature indicating the possible enhancement of student accomplishment in different conditions (Ericsson & Pool, 2016). Lastly, the "Chemers, Hu, and Garcia's model" is a longitudinal study developed by Martin M. Chemers, Li-tze Hu, and Ben F. Garcia at the University of California. They inspected the effects of optimism and academic self-efficacy on students' achievement, commitment to continuing in school, health and stress. Chemers et al. (2001)

These studies show that there is a broad range of research highlighting numerous learning factors affecting students' achievement. Among them, there are many factors outside of students' control, even though they are aware of those factors, they may not be able to address issues associated with factors on their own. For instance, students cannot choose the place where they are born, and they may not be able to change other people's decisions. However, students can control learning factors such as their attitude, psychology, behaviour, self-responsibility skills and most of the cases their physical health. Furthermore, students have the responsibility for their communication and how they want to study and learn. There is a gap in the literature exploring and associating these different perspectives from students' controllable factors (Akama, 2017).

This paper proposes a student-controllable learning factor model that combines the perspectives of Psychology, Self-responsibility, Sociology, Communication, Learning and Health & wellbeing (PS2CLH). The proposed model used qualitative methods to identify underlying aspects affecting academic achievement and selected controllable factors.

Then, data is collected through a self-evaluative web-based questionnaire. Each student's past performance and factors affecting this are then quantified. This study reveals the impact of students' controllable factors on student achievement. The focus of the hypothesised PS2CLH model is on factors which students can control so that students could be aware of how such factors influence their achievements and then take actions to address these issues independently or being taught via mentorship programs.

Proposed PS2CLH'S model

PS2CLH's model is a student-controllable learning factor model for enhancing students' ability to control their performance, which combines the perspectives of Psychology, Self-responsibility, Sociology, Communication, Learning and Health & wellbeing. This research acknowledges the existence of other perspectives, however, the proposed model excludes them for the absence of a scientific study showing the correlation between those perspectives with student's academic performance, such as the religion, spirituality, positive thinking, law of attraction and so on.

The proposed PS2CLH' model, is the big umbrella abstract concept of the PS2CLH's perspectives. Knowing that each perspective has a large spectrum of learning factors affecting students' performance, the application of the model will be adapted to each Country or University's reality. Selecting for each perspective only the most influence learning factors, since each University has their own challenges.

The proposed PS2CLH model was inspired by the area of child development and early learning (Landry, 2014). This field develops children's critical skills through interactive play in a safe and engaging environment. The domains of child development and early learning are categorical organized by: cognitive development; general learning competencies; socioemotional development; and physical development and health [8] (Allen et al., 2015). The categorisation of child development and early learning stems from a variety of sources.

Therefore, there is not a single best categorical organisation. Indeed, it is essential to recognise that the perspectives shown in PS2CLH Fig. 1 are not easily separable. For instance, general cognitive processes also relate to learning competencies, such as persistence and engagement (Allen et al., 2015). Nevertheless, PS2CLH identify the main factors affecting students' achievement, which are on students' daily life control, and recognise that they are interactive and mutually reinforcing rather than hierarchical (Allen et al., 2015). For future students' representation, the diagram merges six perspectives into three pairs.

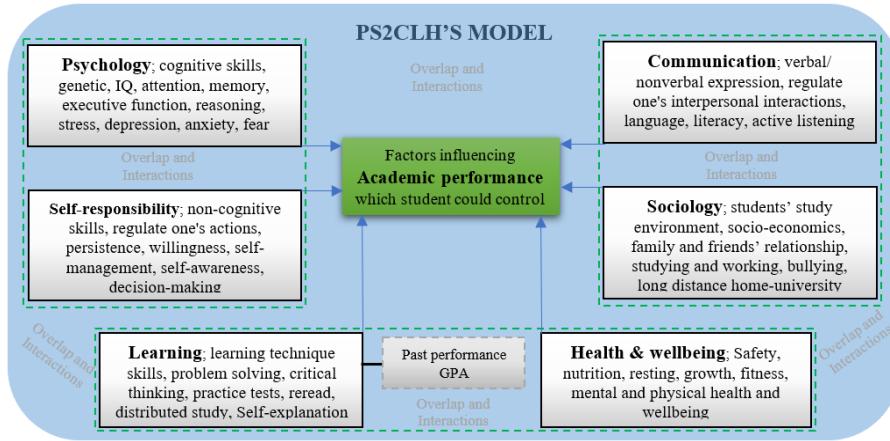


Figure 1: Diagram PS2CLH's perspectives

The “P” on PS2CLH stands for Psychology and “S” for Self-responsibility. These two perspectives relate in terms of internal skills where Psychology represents cognitive skills and Self-responsibility non-cognitive skills. Both perspectives deal with students' internal state where one's mind and the other perspective is related to willingness. Therefore, they are

selected because it is directly related to student performance, and most of the factors affecting students' performance are in the students' control.

Likewise, the diagram presents Sociology "S" and Communication "C", where Social perspective covers the following elements social-interaction, family and friends' relationship and study environment. Communication is essential for human interaction, allowing to express their thoughts and understand others. Excellent communication is vital during the learning process and seems that it could be in the students' control.

The "L" means learning and "H" Health & wellbeing, to a large extent the mental and physical health and wellbeing elements, is reflected in the diagram, and it is the students' control to apply better learning technics and take care of their health & wellbeing. These perspectives direct affect students' performances. There is below, scientific literature to support the proposed PS2CLH's perspectives on the hypothesised model.

A. Psychology and Self-responsibility

According to Cambridge dictionary, Psychology can be defined as logical investigation in which the human behaviour is drifted due to the way the human mind works (Cambridge University Press, 2019). This perspective is important since, students' psychology has a direct impact on their performance. Hattie's list of factors that affect students' academic performance, presents a wide range of studies supporting the idea that psychological learning disabilities, such as autism, dyslexia, ADHD and so on, negatively affect learners' academic performance. On the other hand, considerable research proves that psychological interventions for students with learning needs have a high impact on their achievement.

Self-Responsibility – it can be defined as being capable to acknowledge your own emotional clashes without blaming or projecting upon someone else. The lack or not of this characteristic determines behaviours (Ascensionglossary, 2019). Self-responsibility skills specify behaviours, strategies and attitudes which are thought to underpin victory in life, such as resilience, beliefs, and self-control. The term self-responsibility is associated with non-cognitive skills since it is related to "personal choice and learners' willingness", contrary to genetic or cognitive skill. The concept of 'non-cognitive skill was coined by Bowles and Gintis (Bowles & Gintis, 1976) to emphasise the factors instead of those measured by cognitive test scores. Bowles and Gintis bring to light the role of perseverance, motivation and attitudes instead of academic skills and IQ, as factors contributing to accomplishment. Some studies reinforced their results, confirming the vital role of non-cognitive skills such as student's attitudes, determination and commitment over and above cognitive skill in influencing social behaviour, and students' results and health (Bowles & Gintis, 1976).

The well-known scientist, Michio Kaku, describes the importance of non-cognitive skill. He said that when they look at all the different theories about what makes a successful person. They realise that almost all the theories are wrong because it has been verified, for instance, that high IQ does not determine the outcome or a person's success (Kaku, 2018). He asked what the one psychological test that correlates with success in life is, and he found out that the marshmallows test predicts people's success. The marshmallow experiment created by Walter Mischel, who studied delay gratification in young children, emphasises self-control in human growth. The experience consisted of asking a child if he/she wanted a marshmallow at that moment or two marshmallows an hour from then, and the children that wanted a marshmallow immediately tended to be those who wanted shortcuts, those who did not want to put in the hard work (Mischel, 2015).

Resisting the marshmallow and the success of self-control generated studies such as “Grit” the power of passion and self-control developed by Angela Duckworth (Duckworth, 2017). ‘Grit’ and Self-control are some of the non-cognitive skills which have a strong correlation with students’ performance. However, these skills seem to be more correlated with steady personality rather than soft skill (Lesli & Ingrid, 2013). On the other hand, interventions developed by Wilson in his book title “REDIRECT” (changing the stories we live), his interventions show long term positive outcomes for students (Wilson, 2013).

To summarise psychological and self-responsibility perspectives, the psychology perspective represents the students’ internal state of the mind and behaviour and the self-responsibility is related to the student’s decision. As presented below changing both perspectives will positively impacted on the students’ mind, behaviour and decisions, leading to a better performance. It is moving to the next views, an examination of sociology and communication perspectives.

B. Sociology and Communication

Sociology is usually defined as a social science that studies the human interactions in societies and acknowledge the processes used to change or preserve their way of living (Faris, E.L, 2019). There is a strong correlation of students’ academic results in secondary school with their academic activities, awareness of their adapting study’s strategies, parent’s guidance, family pay, parent’s level of education, and so on. Students’ social facet and their homes plays an important role on students’ performance. The environment at school or the university, for instance, Laiqa et al. (2011, pp.16-707) support the notion that the school facilities affect the education process. Rossi and Montgomery’s model also reinforces this idea where it focuses mainly on social student’s context, which leads to two distinct scenarios. Firstly, the quality of the school climate and curriculum. Second, the quality of home & community environment Chemers et al. (2001, pp.55-64). Laiqa et al. (2011, pp.16-706) introduce an essential factor, which is the fact that the environment directly affects students’ academic performance. Therefore, a home architecture is essential, the shape, colour, texture, scale, proportion and quality of illumination links to the quality of the environment. Those conditions impact on human and cultural behaviour. Although most of the researches focus is at the schools’ level rather than the residence, they bring light to the fact that building conditions associated with personal comfort, which affects students’ performance.

Lawson and Bacolod (Lawson, 2001) support this argument, declaring that the supply of essential services for instance, “the electricity in learning environment improves the concentration of teachers and students” Laiqa et al. (2011, pp. 16-709). Bearing in mind that people’s environment is dynamic and constructed out of family and social relations. Family structure has a direct impact on students’ achievement. Bankston and Caldas (1998, pp.715-723) research states that non-single-headed families are six times more likely to be wealthy than single-headed families. Consequently, students in the single-headed family’s environments will not only have to live without a father or mother figure but also with financial difficulties. Referencing Mulkey et al. (Mulkey et al, 1992, pp.48-65), family structure influences their school performance. Furthermore, Bankston and Caldas (1998, pp. 18-716), reinforce the idea that students’ family structure has its impact on educational success not only socioeconomic status.

Social relationships and Communication play an important role on students' performance. Communication is a process by which messages or specific information is led from a precise spot or individual to another, or the message itself. Likewise, communication can be a trade of information and the manifestation of feeling that culminate in understanding (Cambridge University Press, 2019). According to Noble (Noble et al, 2006), Social communication skills also play an essential role in students' achievement. Therefore, it creates significant upheaval for some students, due to inconsistency between the languages at home and on campus. It created a double-awareness on the process of adapting to a different language environment. On the other hand, according to Abdullah (Abdullah, 2005, pp.1-26) students with efficiency in English and excellent communication skills expand their achievement. Furthermore, William & Burden's (1997) discovery that language in the classroom gives confidence among students to discuss, to use the new terminology to communicate and experiment different forms of conveying meanings as well as to deal with failures and successes.

In brief, the social perspective is related to the student's external environment, and communication is the student's perception of the bridge between external and internal dialogue. Improving students' social relationships and their communication gives to students' a sense of belonging, confidence to express thoughts and facilitate the learning process. Finally, moving on to learning and health & wellbeing perspectives.

C. Learning and Health & wellbeing

Learning could be described as the transformative process of assimilate information after internalizing and blended with we have experienced before, bringing the need to assimilate the new content and expands our experience overall. This process of learning depends on three phases: input, process and reflection (Malamed, C, 2019). There is an outstanding monograph "Improving students' learning with effective learning techniques". Developed at Kent State University in 2013 a study, which presents ten years of literature indicating that they could enhance student accomplishment across a wide range of environments. That study focuses on practical learning techniques a research group led by Dunlosky (2013, pp. 4 –58).

Active learning strategies have been presenting positive results in students' academic performance. Ericsson is considerate to be the research leader of what makes people great in what they do. He coined the term "expertise", inspired the 10.000 hours rule and created the deliberate practice. For 30 years deliberate practice is a powerful method to help children to develop expertise on practical learning. On his recent book, "PEAK" Ericsson calls attention to purposeful practice, divided into four components of purposeful practice. 1. Having a clear goal, 2. Intense Focus, 3. Immediate feedback and 4. Get out of their comfort zone (Ericsson & Pool, 2016).

Wellbeing is characterized as a decent or agreeable state of existence, as well as a state portrayed by happiness, satisfaction, safety and success. (Dictionary.com, 2019). Health & wellbeing is the other perspective that has a significant impact on students' performance. According to academic Singh, there is a correlation between physical activity and scores on various subjects. Ms. Singh argues "There are, first, physiological explanations, like more blood flow, and so more oxygen to the brain. Being physically active means that there are more hormones produced like endorphins. Also, endorphins make the stress level lower and improve the mood, which means better performance" (Singh, 2012). Likewise, students involved in organised sports are more focused on the classroom. However, differences among the observational studies lead Ms. Singh to declare that it is not possible to establish the

correlation between the amount or kind of activity and the level of academic enhancement (Singh, 2012).

In a nutshell, improving learning and health & wellbeing perspective are crucial to rise students' performances, these areas represent the student's strategy and action toward education.

Testing the proposed PS2CLH's perspectives: Angolan context

Angola is a developing Southern African country, which obtained independence in 1975, then lived in a state of civil war until 2002. The war left many psychological and social traumas which affected the Education System directly; as a result, in Angola, the level of students' academic performance is at a considerably low level (UNICEF, 2011, pp 22). In the Country, no scientific studies are pointing on the relationship between factors affecting students' achievement. Therefore, the variable's selection reflects the common belief of what should be the factors that affect the most students' performance. In addition, was taken into consideration the current Angolan reality. Also, the variables selected are self-evaluated variables, that do not need external test.

Data collection was carried out in Luanda capital of Angola, at ICT laboratory of "Universidade Católica de Angola". Data is collected through a self-evaluative web-based questionnaire. Each student's PS2CLH's factors affecting their performance are then quantified using a 5 point Likert scale labelled as 'strongly disagree', 'tend to disagree', 'Do not know' 'tend to agree', 'strongly agree'. Initially, the population sample number was around 500 students from different courses and years, after the clean-up process had around 432 students. The balanced partition split found between results and percentage for the available data was: 60% training, 20% testing and 20% validation. The test model uses the past performance grade point average GPA as the target variable where it is correlated with all other variables selected on the implementation of the PS2CLH model in Angola.

The students range ages was between 18 to 30 years old, both male and female, with a nationality of 98% Angolan, on degree program where the majority of the students was between their second and third years.

Perspectives of Angolan Context PS2CLH's variables

Perspectives	Angolan Context PS2CLH's variables
Psychology	Stress; Feel depressed; anxiety or fear; Disturbance of the mode of being; believe in witchery; low self-esteem; some unspecified psychological problems
Self-responsibility	Set priorities; Establish and achieve personal goals; Time management; Aim for excellence; Procrastination; Immediacy; Accept the change; hours study per day
Sociology	Studying and working; Bullying; Family income; Sensual images; Discouragement and negativity; Long distance; certain negative beliefs and habits; bad conditions of habitability; Lack of electricity; water or sanitation, Lack of public transport; hours play/distraction per day
Communication	Fluency in language; Understand the lecturer in the classroom; Understanding and interpretation of reading; Expressing yourself; Grammar and vocabulary; Stuttering or typology of disfluencies; learning prob which impact my communication
Learning	Preparation of a questionnaire; Highlighting and underlining; Practice tests; Reread; Distributed study; Self-explanation; Prepare summaries; Problem with calculus and mathematic; ADHD
Health & wellbeing	Regular physical activity and exercise; feel mentally healthy; plenty of energy during the study time; eating healthily; rest body and mind; sleep problems.

Table 1. Variables used in Angolan context

Findings and discussion

Among numerous machine learning's models tested, the "RandomForest model", shown at the Appendix session 'Table 2', produced the best result. The prediction result presented by the model, based on proposed PS2CLH's perspectives was 94% of accuracy. Showing that the selected variables from proposed PS2CLH's model directly correlates with the target variable or the student's academic performance. The target variable was the students' past performance, grade point average (GPA). For more detail, find attached the predictor Importance variables at the Appendix session 'Figure 3'. Having "studying and working", one of the most significant predictor importance. Followed by the variables "the average hour students play/distractions per day", "Aim for excellence in everything you do", "Establish and achieve personal goals" and "Practice Tests" and so on. It explains the correlation on time students spend on distractions such as social network, spending hours with friends, and so on. The "Aim for excellence in everything you do" and "Establishing and achieving personal goals" variables also had a significant impact on the model. Dealing well with these last variables has proven to be game-changing for top students, what seems to be in harmony with the previous literature reviewed, such as the marshmallow factor or delay instant gratification and apply deliberate practice.

Students were divided into three groups, the first group 'No Participants' was approximately 50 students that did not fill the questionnaire, the second group was 'Participants, No Interventions', where these students filled the forms but did not receive any interventions, and the third group of students was about 50 students, they filled the questionnaire and had experts' interventions (such as Piagetian programs) on their learning factors. The average score in standard deviations effect size, are shown below.

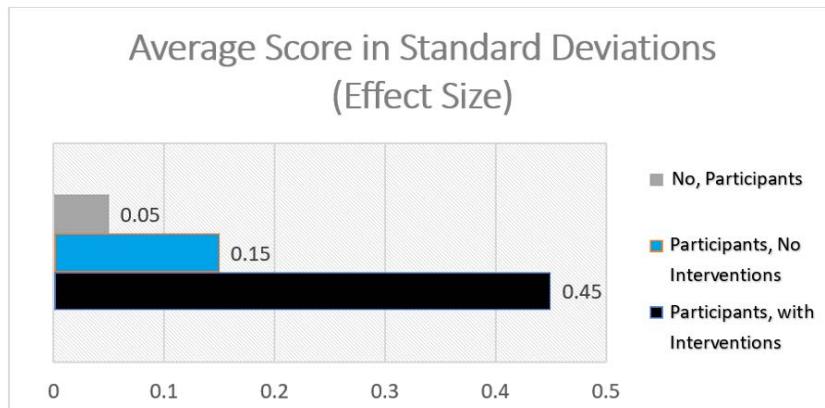


Figure 2: Shows the average score or GPA in standard deviations effect size

Due to lack of human resource, was selected 50 students from the initial 500 students, to work on their major problems presented in the model. At the end of the semester was found that it raised students' awareness of PS2CLH perspectives, helped them manage factors affecting academic performance more effectively.

Consequently, most of the students have enhanced their academic performance by addressing these critical factors. Nevertheless, due to the limitations of the current sample data, the PS2CLH model will be further monitored for various applications.

Conclusion and Future work

The proposed PS2CLH's model presents the psychological perspective, which correlates with the student's internal state of mind. Thus, the self-responsibility standpoint is related to a student's daily decisions. Besides, the social perspective links with students' lifestyle environment. Additionally, a communication standpoint associated with students' perception and how they express themselves; communication could be seen as the bridge between external and internal dialogue. Finally, health & wellbeing and learning perspectives are related to the student's studying strategy and action.

Accordingly, to the results, students raised awareness about PS2CLH's perspectives from factors influencing students' achievement and which are in their control, this helped them to control how much influence the factors have on them. It empowered the students, for instance, students start using a daily task to build new habits.

For future work, a representation of a student's controllable factors can be built, starting by developing a student's web-based questionnaire. Each question would have a weight and the sum of the questions' value in specific coordinate value. Having, psychology & self-responsibility (coordinate/axes X), social & communication (axes Y), learning and health & wellbeing (axes Z). It results in student representation on a point in three-dimensional space 3D. From a mentor's or university manager's point of view, it is essential to measure and to keep track of the student's performance alongside factors which affect their academic achievement, it will be possible to accomplish this intend using 3D student representation.

The PS2CLH's model scalability can be further improved by the development of a Social Network, where each student will have all the problems that most affect them on a personal profile with their own individual learning factors. Thus, each student will be able to independently and proactively work on those issues that affect their results the most.

References

- Abdullah, A. (2005). Some determinants of student performance in Financial Management Introductory course: an empirical investigation. *Journal of King Saudi University Administrative Sciences*, 5(1), pp. 1-26.
- Akama, E., (2017), Student's Performance Prediction Using FP-Tree Data Mining Techniques, *International Journal of Science and Research (IJSR)*, Jnrvu University, Udaipur, India, ISSN: 2319-7064.
- Ascensionglossary. (2019) Meaning of Self-Responsibility <https://ascensionglossary.com/index.php/Self-Responsibility>, Assessed on 25 November of 2019.
- Bankston, CL, Caldas SJ (1998). Family structure, schoolmates, and racial inequalities in school achievement. *Journal of Marriage and the Family*, 60(3): pp. 715-723.
- Bowles, S., & Gintis, H. (1976). *Schooling in capitalist America* (Vol. 75). New York: Basic Books.
- Cambridge University Press, (2019) Definition of communication from the Cambridge Academic Content Dictionary, <https://dictionary.cambridge.org/dictionary/english/communication>, Encyclopædia Britannica, Inc. Assessed on 25 November of 2019.
- Cambridge University Press. (2019) Cambridge Academic Content Dictionary, Psychology <https://dictionary.cambridge.org/dictionary/english/psychology>, Assessed on 25 November of 2019.
- Committee on the Science of Children Birth to Age 8, (2015), “*Deepening and Broadening the Foundation for Success,*” Institute of Medicine; National Research Council; Allen LR, Kelly BB, editors. Washington (DC): National Academies Press (US); Child Development and Early Learning. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK310550/> Assessed on 14 June of 2019.
- Dictionary.com, (2019), Definition of wellbeing, © 2019 Dictionary.com, LLC, <https://www.dictionary.com/browse/wellbeing>. Assessed on 25 November of 2019.
- Duckworth, A., (2017), *Grit: Why passion and resilience are the secrets to success*, Paperback, Scribner, An Imprint of Simon & Schuster, Inc. New York, NY 10020. United States of America, ISBN 978-1-5011-1110-5.
- Dunlosky, J. Rawson, K.A., Marsh, E.J., Nathan, M.J. & Willingham, D.T. (2013). *Improving Students' Learning with Effective Learning Techniques:* Promising Directions from Cognitive and Educational Psychology, aps association for psychological science, Kent State University; Duke University; University of Wisconsin–Madison; and University of Virginia. *Psychological Science in the Public Interest* 14(1) pp. 4 –58.
- Ericsson, A. and Pool, R., (2016), *Peak: Secrets from the New Science of Expertise*, United States of America, New York 10016, ISBN 978-0-5-45623-5.

Faris, E.L, (2019) Meaning of Sociology, <https://www.britannica.com/topic/sociology>, Encyclopædia Britannica, Inc. Assessed on 25 November of 2019.

Lesli, G. M., & Ingrid, S. (2013) *The impact of non-cognitive skills on outcomes for young people, Literature review*, Leading education and social research, Institute of Education, University of London.

John Hattie. (2009) *Visible Learning*, A synthesis of over 800 meta-analyses relating achievement, Routledge, 270 Madison Avenue, New York, NY 10016, ISBN 0-203-887-6.

John Hattie. (2018) *Visible Learning plus 250+ Influences on Student Achievement*, https://www.visiblelearningplus.com/content/research-john-hattie, 250_influences_10.1.2018.pdf, Assessed on 7 December of 2018.

Kaku. M., (2018). *Why You Should Be Optimistic About the Future* | Michio Kaku on Impact Theory. Published by Tom Bilyeu on Apr 17, 2018, site: <https://www.youtube.com/watch?v=tGulK44YaOM>, Assessed on December 10 of 2018.

Laiqa, R., Shah, R. U. Khan, S. M. (2011). Impact of quality space on students' academic achievement, *International Journal of Academic Research*, pp.706-711.

Landry, S. H. (2014), *The role of parents in early childhood learning*, Children's Learning Institute; University of Texas Health Science Center, USA.

Lawson, B. R (2001). *The Language of Space*. Boston: The Architectural Press.

Malamed, C. (2019) From The New Social Learning by Tony Bingham and Marcia Conner, <http://theelearningcoach.com/learning/10-definitions-learning/>. Assessed on 25 November of 2019.

Martin M. Chemers, Li-tze Hu, and Ben F. Garcia, (2001), Academic Self-Efficacy and First-Year College Student Performance and Adjustment, *Journal of Educational Psychology* Vol. 93, No. 1, 55-64.

Mischel. W., (2015). *The Marshmallow Test: Understanding Self-control and How To Master It*, Paperback London, Penguin Random House UK, ISBN 9780552168861.

Mulkey, L. M. Crain, R. L, Harrington, A. J (1992). *One-parent households and achievement*: economic and behavioural explanations of a small effect. *Sociology of Education*, 65: pp. 48-65.

Noble, J. P., Roberts, W. L., & Sawyer, R. L. (2006). *Student Achievement, Behavior, Perceptions, and Other Factors Affecting ACT Scores*. ACT Research Report Series, 2006-1. ACT, Inc.

Singh, A., (2012), *Does Physical Activity Lead to Higher Grades?* Education Report, VU University Medical Center, Amsterdam.

United Nations Children's Fund, UNICEF. (2011). Progress Evaluation of UNICEF's *Education in Emergencies and Post-Crisis Transition*, Programme: Angola Case Study, Evaluation Report, pp. 22.

Williams, M., & Burden, R. L. (1997). *Psychology for language teachers*: a social constructivist view. New York: Cambridge.

Wilson. T., (2013). *Redirect: Changing the Stories We Live By*. Paperback, Penguin Group, London, England, ISBN 978-0-141-04224-4.

Contact email: ara0223@my.londonmet.ac.uk

Appendix

We have applied the SMOTE function to balance the lack of positive values “1 = 65”, “0 = 367”. Then applied the “Random Forest” function to build the model

Having Created the model, we tested and validated, the accuracy was $\approx 0.943\%$

```
0   1  
367 65  
> TrainSet <- SMOTE(Astudent ~ ., TrainSet, perc.over = 500, perc.under=120)
```

```
> table(TrainSet$Astudent)
```

```
0   1  
390 390
```

Call:

```
randomForest(formula = Astudent ~ ., data = TrainSet, importance = TRUE)
```

Type of random forest: classification

Number of trees: 500

No. of variables tried at each split: 8

OOB estimate of error rate: 4.54%

Confusion matrix:

```
0 1 class.error  
0 370 20 0.05128205  
1 31 359 0.07948718
```

Call:

```
randomForest(formula = Astudent ~ ., data = TrainSet, ntree = 500, mtry = 8, importance = TRUE)
```

Type of random forest: classification

Number of trees: 500

No. of variables tried at each split: 8

OOB estimate of error rate: 4.92%

Confusion matrix:

```
0 1 class.error  
0 368 22 0.05641026  
1 32 358 0.08205128
```

```
> table(predTrain, TrainSet$Astudent)
```

```
predTrain 0 1  
0 390 0  
1 0 390
```

```
> # Predicting on Validation set
```

```
> predValid <- predict(model1, ValidSet, type = "class")
```

```
> # Checking classification accuracy
```

```
> mean(predValid == ValidSet$Astudent)
```

[1] 0.9425926

```
> table(predValid, ValidSet$Astudent)
```

```
predValid 0 1  
0 94 2  
1 2 12
```

predict_unseen

```
0 1  
0 130 7  
1 5 19
```

Table 2. Application of the Random Forest function

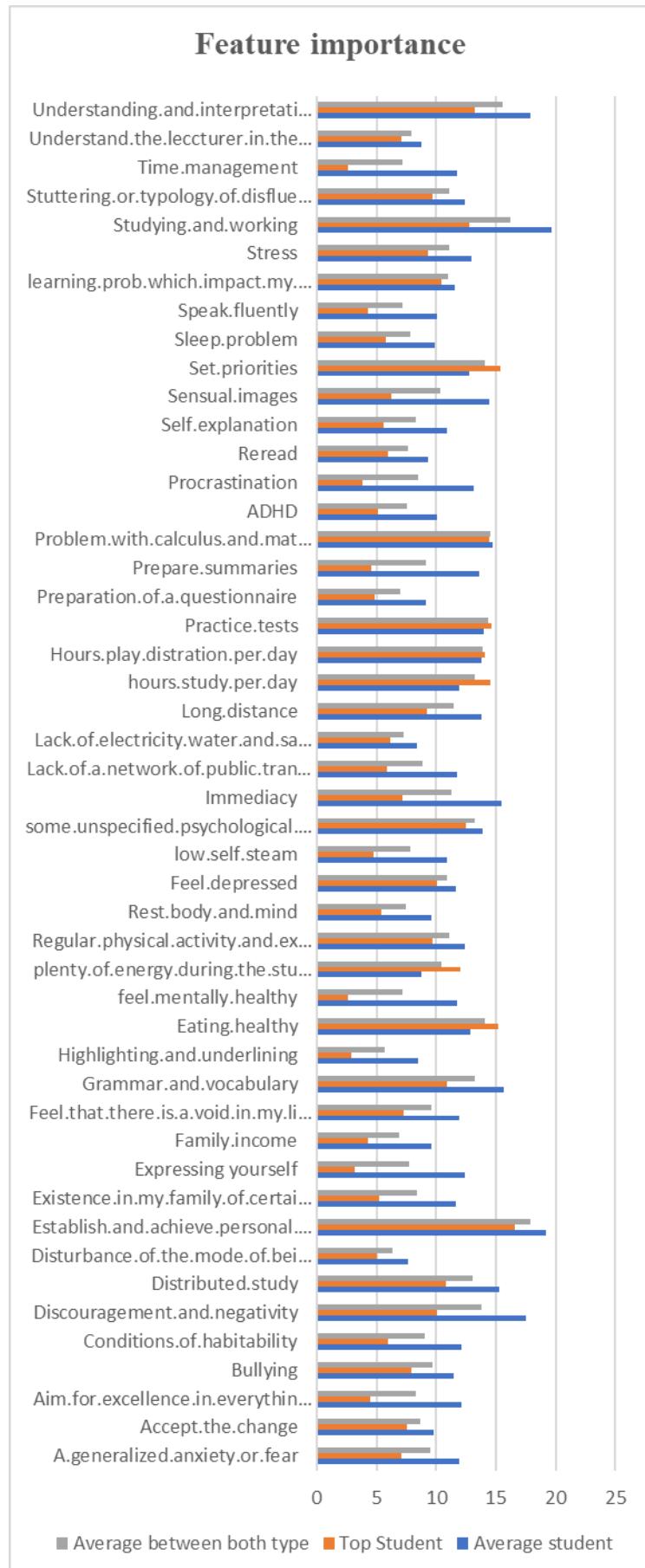


Figure 3: Feature importance Angolan context