

# Leveraging Non-Cognitive Student Self-Reports to Predict Learning Outcomes

Kaška Porayska-Pomsta<sup>1</sup>, Manolis Mavrikis<sup>1</sup>, Mutlu Cukurova<sup>1</sup>, Maria Margeti<sup>2</sup>, and Tej Samani<sup>2</sup>

<sup>1</sup> UCL Knowledge Lab, University College London

<sup>2</sup> Performance Learning Education Ltd.

**Abstract.** Metacognitive competencies related to cognitive tasks have been shown to be a powerful predictor of learning. However, considerably less is known about the relationship between student’s metacognition related to non-cognitive dimensions, such as their affect or lifestyles, and academic performance. This paper presents a preliminary analysis of data gathered by Performance Learning Education (*PL*), with respect to students’ self-reports on non-cognitive dimensions as possible predictors of their academic outcomes. The results point to the predictive potential of such self-reports, to the importance of students exercising their self-understanding during learning, and to the potentially critical role of incorporating such student’s self-reports in learner modelling.

## 1 Introduction

Academic performance is typically measured through assessments on standardised tests, which are often used formatively together with teachers’ assessments of students’ attitudes and motivation to predict final grades on high-stake exams. However, there is no standardised way in which the predicted grades do actually reflect students’ attitudinal and motivational traits. Although the importance of students’ emotions, motivation and lifestyles to learning is confirmed by substantial research [1], the lack of teacher training with respect to ‘diagnosing’ students motivation and attitudes, coupled with the known challenges related to accessing reliably other people’s mental states (e.g. [2]), leads to a whole variety of subjective judgements which are hard to validate and operationalise in everyday educational practices.

Owing to a prevalent emphasis on subject-specific education together with a predominantly didactic, ‘*teacher in charge of the assessments*’ approaches that are adopted in mainstream education, one aspect which is often overlooked in school contexts is the value of engaging students in their own assessments, for example through self-reporting on how they feel, what motivates them and what they consider the possible barriers to their learning. Such self-assessments, involving the metacognitive competencies of self-monitoring, play a crucial role in allowing students to reflect on their experiences, motivation and attitudes and in helping them first, to understand their own behaviours, and second, to plan actions, set goals and aspirations, and ultimately – to make informed decisions [3].

Although metacognitive competencies have been established as a powerful predictor of learning [4], they have been mainly examined in literature and applied in educational practice in relation to subject specific tasks, e.g. in explaining away problem solving [5]. Self-explanations by students in relation to socio-emotional, attitudinal and life-style choices do not figure explicitly on the mainstream curriculum agenda, despite the observed benefits, because: (1) it is not clear to what extent students' self-explanations in relation to such non-cognitive dimensions are actually reflected in their academic performance; (2) asking students to self-explain in terms of such dimensions may be considered intrusive, especially in whole classroom situations. AIED approaches, especially those concerned with learner modelling and open learner models, can potentially address both those concerns and provide a powerful means for supporting students in engaging in self-monitoring and in self-regulation. In this paper we present a preliminary analysis of data generated by Performance Learning Education (PL) in live high-stakes intervention contexts in two schools (A and B), involving a total of 48 students aged 16-17 preparing for pre-university examinations. We describe the self-reporting instrument (PLOA) used to elicit responses from the students on key aspects of their socio-emotional states and lifestyle habits related to a potential risk of their underachieving in their exams.

## **2 Performance Learning Online Assessment: PLOA**

PL supports schools in raising students' academic achievement as measured through high-stakes exams, which are required for entry to university, further educational, or professional development. The company works at all levels of education from primary and secondary, through further education to the university level. Its main user base are state-funded schools catering for pupil cohorts from low socio-economic backgrounds. The attainment in those schools tends to be poor and often below national average, with many pupils leading hectic lives and not prioritising or valuing academic achievement.

PL's approach emphasises the importance of students developing a good understanding of their own strengths and weaknesses along with the possible causes for both. Its approach is structured around four overlapping trait categories known to be of critical importance to students' learning and development: (i) *motivation*, linked to student's goal orientation; (ii) *organisation* related to the executive functions of planning and attentional control; (iii) *memory* broadly related to attentional control and cognitive flexibility; (iv) *lifestyle* related to sleep and physical and emotional wellbeing of the students. There are two elements to PL: (1) self-assessment by students aimed to ascertain their strengths and weaknesses along the four trait categories described; (2) curriculum, which coaches students in how to attend to specific aspects of their lifestyles, attitudes, emotions and goal management.

The students' self-assessments (henceforth referred to as PLOA) are conducted online at the start and end of pupils' PL's curriculum, to establish

any changes in their self-assessments over time. Twenty eight questions are used to elicit self-assessment from individual students in relation to the four categories assessed. Student's responses are scaffolded through multiple choice questions that are associated with each trait category, each question linked to a risk level (1: low risk, to 5: high risk), with each level representing the relative degree to which a particular trait may be a barrier to a given student's academic achievement. For example, if under the motivation category a student declares that they cannot cope with and tend to panic under pressure, this is linked to a relatively high PLOA risk level. PLOA scores associated with each student answer choice are aggregated at the end of the assessment and an overall PLOA is calculated using a PL's proprietary weighted means function which is further associated with percentage ranges; the lower the percentage, the higher the PLOA risk level. For further details about PL's approach and the development thereof see [6].

### 3 Data analysis and Results

Two UK schools which use PL as an approach to raise student attainment have been included in the analysis presented. Both are co-educational secondary schools, catering for between 400 (School A) and 1300 (School B) pupils aged 11-18 years old. In both schools there is an equal boy-girl ratio and both schools have a medium to high free school meals percentage which is used as a key socio-economic school indicator in the UK. Additionally, school A has a special educational needs provision for pupils with moderate learning difficulties.

Initial and final PLOA, gathered before and after PL intervention, were used in the analysis, along with the estimated exam grades (EEGs) and final exam grades. Both sets of grades were provided by the schools and were based on subject-specific tests and, in case of EEGs, on test results and assessments of individual students' attitude and overall effort in each subject. For both schools Pearson correlations coefficient analysis revealed significant correlations between students' PLOA and their final overall grades (School A  $N = 35$ ,  $r = .583$ ,  $p < .001$ ); (School B  $N = 13$ ,  $r = .878$ ,  $p < .001$ ). Partial correlations analysis for each subject for which PL intervention was given, revealed a similar pattern. In addition, a paired-samples t-test was conducted to test the change in the initial and final PLOA. As expected, this significant increase was also apparent in each school. School A: Initial PLOA ( $M = 48.522$ ,  $SD = 10.293$ ) vs. final PLOA ( $M = 53.456$ ,  $SD = 11.617$ )  $t(34) = 12.646$ ,  $p < 0.00$ ; School B: Initial PLOA ( $M = 56.269$ ,  $SD = 10.934$ ) vs. the final PLOA ( $M = 56.654$ ,  $SD = 11.998$ ),  $t(12) = 7.028$ ,  $p < 0.00$ .

As part of a regression analysis, we explored the respective potential of (i) the EEGs and (ii) EEGs together with PLOA in predicting the final exam grades. The results show that EEGs alone explain 26.1% of the variance of the average final score, whereas EEGs with PLOA explain 29.7% of the variance. This statistically significant increase in variance suggests that using both predictors as part of a regression model is more accurate when EEGs and PLOA are used together. The results are summarised in Table 1.

**Table 1.** Regression coefficients and statistics

	School A (N=36)			School B (N=13)		
Predictors	Coefficients t		Sig	Coefficients t		Sig.
Constant	-1.97	-1.96	.058	-1.614	-1.661	.128
Overall initial score	.504	4.341	.000	.181	1.147	.278
Final PLOA score	.489	4.217	.000	.800	5.069	.000
Overall model	$r = .764, r^2 = .584, p < .00$			$r = .893, r^2 = .798, p < .00$		

To help interpret these results, and given that the long term goal of this research is to automate the process of predicting student learning outcomes (here as measured by exam grades) in order to offer personalised support to different students, we also aggregated the data across schools and subjects. As expected, there is a significant increase in the PLOA for the aggregated data set  $t(47) = 7.028, p = p < .00$ . The corresponding regression analysis also resulted in a significant model,  $F(2, 45) = 19.383, p < .001, r = .680, R^2 = .463$  with both EEGs and PLOA being significant predictors ( $score_f = -1.166 + .617 * score_i + .053 * ploa_f$ , standardized Beta  $score_i = .416, ploa_f = .459, p < .000$ ).

## 4 Discussion and Conclusions

Whilst the main limitation of the analysis presented is that it is based on small number of students, the fact that the PLOA increases significantly suggests that the PL's curriculum, which coaches students in how to attend to specific aspects of their lifestyles, attitudes, emotions and goal management is effective. The regression analysis suggests that final PLOA together with EEGs predict the final scores better than either one alone. To contextualise this, the regression function together with the corresponding coefficients indicate that as the average initial score increases the average final score will increase by around 62% (if the final PLOA is held constant). Similarly, as the final PLOA increases by one unit (one percentage), the average final score will increase by 5% (if the initial score is held constant). Considering that these results were generated in low-achieving schools in which the students taking part in the PL intervention are the lowest of the achievers, together with the fact that the intervention was of a relatively short duration (a total of eight hours per school), provides a particularly powerful motivation for investing further in understanding and modelling the relationship between students' self-explanations and metacognitive competencies, especially as relates to students' attitudes, motivation and lifestyles, and their academic achievement in high-stakes exams. The results also suggest that providing a systematic and consistent way in which student's can self-report on those 'non-cognitive' aspects of their learning may be an important learning and assessment tool.

In summary, the results of the analysis presented in this paper are promising with respect to supporting the long-term goal of this research to develop and automate further the student self-reporting functionality of the PL approach.

## References

1. Immordino-Yang, M.H., Damasio, A.: We feel, therefore we learn: The relevance of affective and social neuroscience to education. *Mind, Brain, and Education* **1**(1) (2007) 3–10
2. Porayska-Pomsta, K., Mavrikis, M., S., D., Conati, C., R., B.: Knowledge elicitation methods for affect modelling in education. *International Journal of Artificial Intelligence in Education* **22**(3) (2013) 107–140
3. Terricone, P.: The taxonomy of metacognition. (2011)
4. Richardson, M., Abraham, C., Bond, R.: Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological bulletin* **138**(2) (2012) 353
5. Alevan, V.A., Koedinger, K.R.: An effective metacognitive strategy: Learning by doing and explaining with a computer-based cognitive tutor. *Cognitive science* **26**(2) (2002) 147–179
6. Samani, T., Porayska-Pomsta, K., Luckin, R.: Bridging the gap between high and low performing pupils through performance learning online analysis and curricula. In: *Proceedings of the International Conference on Artificial intelligence in Education. AIED'17*, Springer-Verlag (2017) 650–655