

Does Self-Discipline impact students' knowledge and learning?

Yue Gong, Dovan Rai, Joseph E. Beck and Neil T. Heffernan
Computer Science Department, Worcester Polytechnic Institute

Abstract. In this study, we are interested to see the impact of self-discipline on students' knowledge and learning. Self-discipline can influence both learning rate as well as knowledge accumulation over time. We used a Knowledge Tracing (KT) model to make inferences about students' knowledge and learning. Based on a widely used questionnaire, we measured students' level of self-discipline. When we analyzed the relation of students' self-discipline with their knowledge attributes, we found that high self-discipline students had significantly higher initial knowledge, but there is no consistent relationship of learning while using the tutor. Moreover, higher self-discipline students seemed more careful with respect to making careless mistakes.

1 Introduction

Intellectual attributes (e.g., long term memory, ability to think abstractly) and nonintellectual attributes (e.g., motivation, self-discipline) both contribute to a student's academic performance [1]. Intelligent Tutoring Systems (ITS) have focused on cognitive aspects over last 25 years and are now becoming increasingly aware of non-cognitive traits like motivation, engagement, flow etc. [5,6]. However, self-discipline is still not a major area of exploration in ITS though it has been one of the key areas in psychology and sociology [8,9]. Given that a lot of such large-scale psychosocial studies have been able to demonstrate a positive relation of self-discipline with performance and achievement, we were interested in two questions:

- Does self-discipline have a significant impact when it comes to knowledge acquisition within ITS?
- Does the ITS community need to consider self-discipline while designing ITS?

In this paper, we are trying to use educational data mining technique with fine grained models to get a crisper look at the impact of self-discipline on students' cognitive aspects. We used Knowledge tracing (KT) [3], an established approach to model student knowledge. We can observe students' performance in an ITS over a period of time and make inferences about their latent characteristics like knowledge level and learning across the time. Once we detect those attributes, we can see the impact of self-discipline on the immediate learning and prior accumulation. Besides learning, self-discipline can influence other attributes like consistency and carefulness that can improve performance given the same content knowledge.

2 Methodology

For this study, we used data from ASSISTment, a web-based math tutoring system. We used the data from 171 twelve- through fourteen-year old 8th grade students in urban school districts of the Northeast United States. These data consisted of 74,394 log records of ASSISTment during the period Jan 2009-Feb 2009. We recorded performance records of each student across time slices for 106 skills (e.g. area of polygons, Venn diagram, division, etc).

2.1 Measuring self-discipline

For exploring student individual differences in self-discipline, we employed a questionnaire survey, Brief Self-Control Scale (BSCS; [9]) in December 2008 before the students used the tutor. BSCS is a 13-item questionnaire to measure self-regulatory behavior in four domains: thoughts, emotions, impulses, and performance.

Each question (e.g. “I am lazy”, “I am good at resisting temptation”) asks the respondent to choose from a 5-point Likert scale answer list: a. Very much like me, b. Mostly like me, c. Somewhat like me, d. A little like me, e. Not like me at all. We assigned each response -2, -1, 0, +1, +2 points respectively. We dropped an original survey question (“I wish I had more self-discipline.”) as we find difficult to interpret whether agreeing this statement would imply high self-discipline or low.

2.2 Lie test

While using self-report measures, we have no way of ensuring that respondents don't lie or answer haphazardly. Therefore, we created three criteria to detect lies and out of total 171 students we dropped 31 from our analyses.

1. The questionnaire asked students for their gender. Twelve students gave an incorrect response. Suspecting them not being serious about the survey, we excluded those students from our study.
2. Some students might be randomly picking answers and therefore we checked for consistency in their answers. Among 12 questions in the survey, for 8 of them “Very much like me” implies low self-discipline (e.g. “I have a hard time breaking bad habits”), and for 4 of them, “Very much like me” implies high self-discipline (e.g. “I am good at resisting temptation”). For both types of questions we used the scoring system in Section 2.1. If a student answered “Very much like me” for a question of the first type, he received -2 points. If he answered “Not like me at all” for a question of the second type, he received +2 points. The two responses consistently tell that he has low self-discipline. The sum is zero. But if he had answered “Very much like me” in the second type, the answers are not consistent and the sum of responses is -4. Similarly, if he had answered “Not like me at all” in both questions, that would be still inconsistent and sum would be +4.
3. For each student, we took average of points in both types of questions (since the groups are of unequal size) and summed the two averages and calculated the absolute value. The sum value can range from 0 (completely consistent) to 4 (completely inconsistent). Based on the questionnaire composition and distribution of the sum from our data, we found 1.6 to be a reasonable cut point and dropped 11 students with sum greater than 1.6.
4. We selected two pairs of questions which are basically asking same trait in opposite ways. For example “I do certain things that are bad for me, if they are fun” and “I refuse things that are bad for me” state the same trait. We dropped students who are saying “very much like me”/ “Mostly like me” or “not like me at all” in both questions. There were 19 such students among which 5 were already excluded from step 2.

Finally, our dataset narrowed down to 134 students, with their 68285 log records. We excluded 10% of the records and 20% of the students. For each student, we had 12-dimensional vectors representing their responses corresponding to each survey question. We performed a factor analysis to reduce data dimensions and we used the strongest factor as the student’s self-discipline score.

2.3 Knowledge tracing model

We used knowledge tracing in Dynamic Bayesian Networks (DBN), see Figure 1, to make inferences about student knowledge based on his performance.

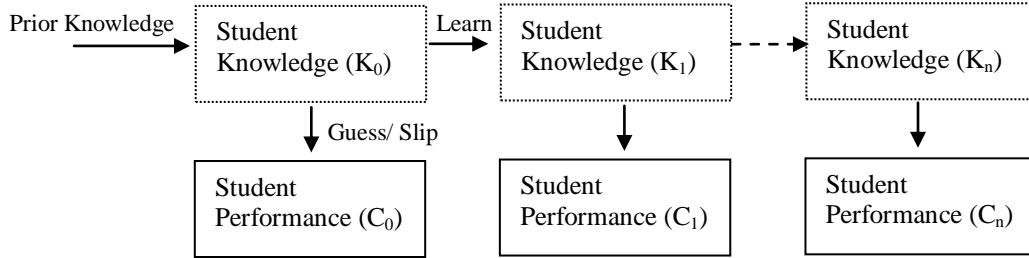


Figure 1. Knowledge tracing model: Dynamic Bayesian network

Student performance is assumed to be a noisy reflection of student knowledge, mediated by two performance parameters guess and slip. The guess parameter represents the fact that the student may sometimes generate a correct response in spite of not knowing the correct skill. For example, some ASSISTment items are multiple choice, so even a student with no understanding of the question could generate a correct response for those. The slip parameter acknowledges that even students who understand a skill can make an occasional careless mistake [3]. The learning rate parameter estimates the probability that student learns new knowledge that he has not known before.

$$\text{Prior Knowledge} = Pr(K_0 = \text{True})$$

$$\text{Guess} = Pr(C_n = \text{True} \mid K_n = \text{False})$$

$$\text{Slip} = Pr(C_n = \text{False} \mid K_n = \text{True})$$

$$\text{Learning rate} = Pr(K_n = \text{True} \mid K_{n-1} = \text{False})$$

We used Bayes Net Toolkit for Student Modeling (BNT-SM [4]), which inputs data and a compact XML specification of a Bayes net model to describe causal relationships among student knowledge and observed behavior. BNT-SM gives us knowledge parameters, prior knowledge and learning as well as performance parameters, guess and slip.

3 Results

3.1 Knowledge tracing model per skill

Based on self-discipline score, we divided students into three equal-sized groups having relatively high, medium and low self-discipline level. For each subgroup, we trained separate knowledge tracing models, and thus estimated knowledge and performance parameters that corresponded to each group. We trained a knowledge tracing model for each of the 106 skills. I.e. observe all the training data across all students for each skill

and derive a set of parameters (Prior knowledge, learning, guess, slip) for each skill. Then, for each self-discipline subgroup, we calculated the median values across all the skills (see Table 1). We report median rather than mean values to avoid unnecessarily weighting outliers. However, in accordance with standard convention, our statistical analyses are based on the means rather than medians.

Table 1: Knowledge and performance parameters for self-discipline groups

	High		Medium		Low	
	Value	P-value (vs. medium)	Value	P-value (vs. Low)	Value	P-value (vs. High)
Prior Knowledge	0.56	4.96E-8	0.48	0.45	0.49	3.62E-6
Learning	0.13	5.68E-6	0.17	0.001	0.14	0.186
Guess	0.38	0.015	0.36	2.38E-6	0.32	1.72E-8
Slip	0.16	8.37E-18	0.20	0.591	0.21	1.09E-18

From Table 1, we see that for prior knowledge, the high self-discipline students are statistically higher than the medium group, and the medium and low groups are statistically tied. Meanwhile, high self-discipline students made more correct guesses and fewer slips relative to their lower self-discipline peers. A higher guess parameter should not be viewed as a bad thing. Consider that guess means the ability to answer a question despite not having mastered the skill. Consider two students with similar partial knowledge and one takes more care to figure the right answer while other quickly asks for help. The model will treat this as a guess by the first student. Such behavior seems related to self-discipline. Similarly, students who are more careful and detail-oriented will make fewer slips (keep in mind that a “slip” is defined as making a mistake in spite of the skill being known). The result shows that higher self-discipline students have more prior knowledge and they are more concerned and careful on their task.

However, we received an inconsistent pattern in the learning parameter. The learning rate of the medium self-discipline group is higher than both the high and low groups. We were concerned with the possibility of overestimating the learning parameter in the medium group by giving the guess parameter less weight, while underestimating it in the high group by giving guess more weight. This concern is due to problems with estimating knowledge tracing parameters [8]. For example, a high “guess” parameter can result in students performing well, but allegedly having little knowledge. Since student knowledge is not directly observable, it is hard to validate the parameter estimates and we are left trusting our model that two groups could perform equally well but one group knows less (see [8] for a fuller discussion of the problems of underdetermined models). To guard against this concern, we also plotted student performance as a function of practice opportunity so that we can see the cumulative effect of the knowledge and performance parameters in students’ future performance for each level of self-discipline.

By using the four parameters of each subgroup and the knowledge tracing equations listed below, we computed the theoretical performance curves for each of them. Specifically, we initialize *knowledge* to be K_0 . After each practice opportunity, we update

knowledge in formula I (below) as the new likelihood of the student knows the skill after the previous practice. Also we compute *performance*, the probability of the student will respond correctly in the current practice opportunity, by using formula II to combine the estimated knowledge with the slip and guess parameters. Intuitively, the probability of making correct response is dependent on student’s knowledge given that he does not slip and also on his probability to make right guess in absence of the knowledge.

I: $Knowledge = previous\ knowledge + (1 - previous\ knowledge) * learning\ rate$

II: $Performance = knowledge * (1 - slip) + (1 - knowledge) * guess$

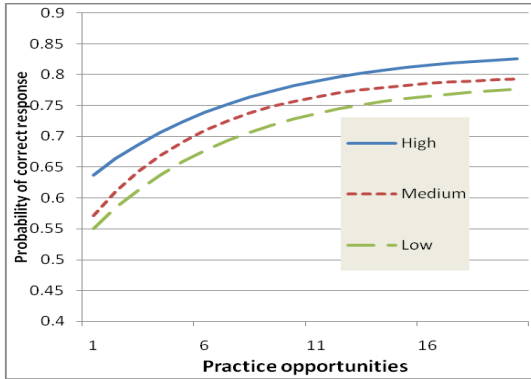


Figure 2a. Theoretic performance curve of three self-discipline groups

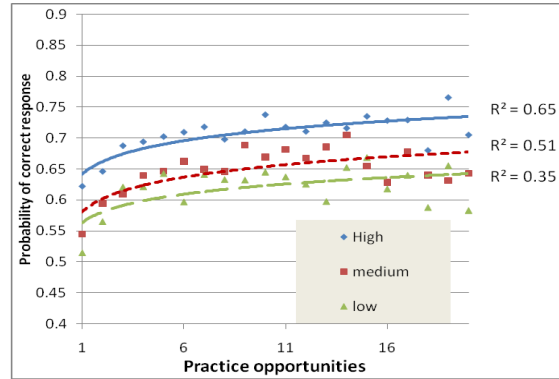


Figure 2b. Real performance curve of three self-discipline groups

From the performance curve in Figure 2a, we see that the combined effects of the four model parameters result in higher self-discipline students performing better. The real performance curve in Figure 2b also showed a similar trend. One interesting observation is to examine the best-fit power curves for each group. The high self-discipline students are performing more lawfully (i.e. higher R^2) than those with low self-discipline, suggesting students with higher self discipline are more consistent. Simply looking at the learning parameters does not tell the whole story. High group students might be learning slower but they are better able to use their partial knowledge to perform better—at least that is what our model is suggesting.

Based on all these findings, we built a causal model that unifies cognitive and non-cognitive aspects of our students. While knowledge parameters like prior knowledge and learning are cognitive attributes, the performance parameters, guess and slip are more related to non-cognitive attributes. This model accounts for the results in Table 1, and suggests the performance parameters might be an interesting avenue of research in their own right (typically the knowledge parameters are of more interest).

3.2 Knowledge tracing model per student

While training a KT model per skill is the regular approach, it is also possible to instead train one model per student by observing his responses in all questions across skills. The model then estimates a set of parameters (prior knowledge, guess, slip and learning) for each student which represents his aggregate performance across all skills. We then

looked for a relationship between the student’s self-discipline score and his knowledge parameters (prior knowledge and learning). As seen in Table 2, self-discipline is positively correlated with student’s prior knowledge (K_0), but again there is no statistically reliable correlation with the learning parameter. In the other words, students with higher self-discipline have more incoming knowledge than their lower self-discipline classmates. However, self-discipline seems not to contribute student’s ability to learn more in each learning opportunity within the tutor. Perhaps higher self-discipline results in having more learning opportunities rather than learning more from each one?

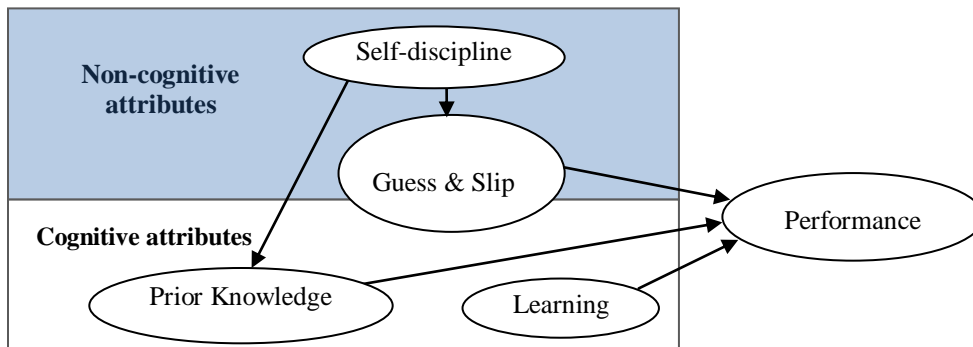


Figure 3: Causal model of cognitive and non-cognitive attributes for academic performance

Table 2: Correlation of self-discipline and knowledge parameters

Correlation of self-discipline with	Correlation	P-value	N
Prior knowledge (K_0)	0.29	0.001	134
Learning	0.13	0.127	134

We also found an interesting observation that self-discipline is highly correlated with the number of problems solved. We were then confronted with two possibilities: either higher self-discipline students are more on task and solve more problems, or students with higher self-discipline have higher knowledge and so need less help and solve problems more quickly. When we did partial correlation within these three variables, as seen in Table 3, we found evidence for the latter possibility. Once we account for prior knowledge, there is no relationship between self-discipline and number of problems solved.

We built a causal model, Figure 4, based on the finding that the higher self-discipline students in fact solved more problems as they were equipped with more knowledge and, perhaps surprisingly, not because they were on task more. The direct correlation between self-discipline and knowledge is 0.29, and between knowledge and number of problems solved is of 0.55. The partial correlations are more interesting. The partial correlation of self-discipline and number of problems, partialing out knowledge was only 0.11. Thus, there is not a direct relation between the two. The partial correlation of knowledge and number of problems solved, partialing out self-discipline is 0.52, i.e. the correlation is relatively unaffected. Thus, knowledge appears to be the direct causal link for number of problems solved, and self-discipline is causally upstream of knowledge.

Table 3: Partial correlation of self-discipline, prior knowledge and # of problems solved

	Correlation	p-value
# of problems solved vs. Self-discipline (prior knowledge as control)	0.11	.22
# of problems solved vs. prior knowledge (Self-discipline as control)	0.52	.000

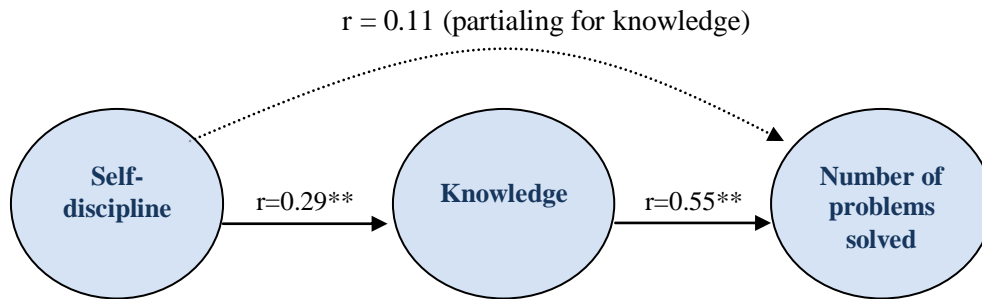


Figure 4: Causal model of self-discipline, knowledge and number of problems solved

3.3 Model validation

KT model parameters can be sensitive to erroneous factors like wrong priors, insufficient data, etc. Therefore, we were curious to try some validation of our model parameters with external data.

To validate our model, we used results from a pretest and posttest on the same set of students. The pretest consisted of a 33-item algebra quiz on the subset of knowledge components that we are using in our models. After a month, the students were presented with posttest with exactly the same questions as in the pre-test. The pretest was performed when the students started using the tutor, and the student's score is used to indicate the amount of incoming knowledge before using ASSISTment. Therefore it works as a standard against which to validate student prior knowledge (K_0) that we estimated in our models.

Also, we calculated students' estimated performance after 8 practice opportunities (P_8) as they practice 8 times on average for each skill during the one month period. We estimate performance from prior knowledge, learn, guess and slip parameters as given by knowledge and performance equations mentioned in 3.1. The correlation of P_8 with posttest can be a measure of validation of the other three parameters.

There is strong positive correlation between the student pretest scores and model's estimation of their prior knowledge. P_8 and posttest scores are also reliably correlated even when we partial out initial knowledge (K_0). I.e. our performance measure is capturing student learning, not just the student's overall level of knowledge. In addition, Figure 5 shows student learning between pretest and posttest.

The gains in Figure 5 are consistent with our KT model results. The high self-discipline group has higher incoming knowledge than both groups and their final performance is also highest. But when it comes to learning, the medium group appears to have the

highest gain. So, we considered some possibilities for the explanation of lower learning in high group. One reason for lower learning in high group could be due to the fact that they already have high knowledge and it is harder to have more gain when we start from higher value. For example going from 50 to 60 is easier than going from 80 to 90.

Table 4: Correlation of prior knowledge (K_0) and P_8 vs. pre and post-test respectively

	Correlation	p-value	N
K_0 vs. Pretest	0.80	3.80E-31	134
P_8 vs. Posttest	0.77	2.57E-25	123
P_8 vs. Posttest (partialing out K_0)	0.34	1.58E-15	123

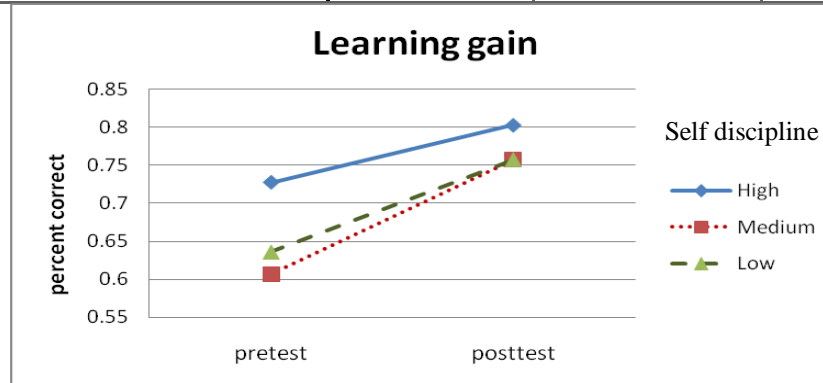


Figure 5: Comparison of learning gain

To test this possibility, we divided pre-test scores into three bins and performed an ANOVA. We treated pretest and self-discipline as factors in our model since we did not necessarily expect a linear effect (as would be implied by treating them as covariates). Table 5 shows the estimated marginal means of gain score for each level of the factors.

Table 5: ANOVA analysis gains by pretest and self-discipline

Pretest	Self-discipline	Mean gain	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
>80%	High	0.03	0.03	-0.03	0.10
	medium	0.06	0.07	-0.07	0.19
	low	-0.03	0.05	-0.14	0.08
(40-80%)	high	0.07	0.03	0.02	0.13
	medium	0.15	0.02	0.10	0.20
	low	0.12	0.03	0.07	0.18
<40%	high	0.15	0.05	0.04	0.25
	medium	0.25	0.05	0.16	0.34
	low	0.05	0.04	-0.03	0.13

From the result in Table 5, we see that medium group has higher learning in all bins (i.e. they are learning faster no matter what their starting level in pre-test is). Therefore, it

appears that the medium group indeed has higher learning and maybe having a balance of self-discipline and some spontaneity helps in having better learning gains. However, their lower incoming knowledge makes the idea of a higher learning rate difficult to reconcile. We choose to leave this as an open discussion for further experimentations in future.

4 Contributions

Psychosocial studies have been based on performance measures like report cards, GPA, income, college admission, etc. [1,8,9]. But, our fine grained model gives us the tools to measure their performance and also latent attributes like knowledge, learning, and even guess and slip. We have found that the impact of self-discipline on students using computers is complex, and appears to influence knowledge and performance while using the tutor. We have constructed a causal model of the impacts of cognitive and non-cognitive attributes on performance within an ITS, and showed that the variability in performance is not only dependent upon cognitive attributes, but also on other non-cognitive aspects like carefulness and self-discipline.

We modified the regular approach to train KT model with data per skill and instead estimated per-student parameters. Although a per-student model trained on prior users is not useful to ITS designers (since it does not apply to new students using the system!), performing parameter estimation at the individual level can open new ways to make different analyses with other individual characteristics. With this new approach, we were able to make correlations of students' pre-test and post test with their knowledge and performance estimations, thus validating the model parameters.

5 Future work and conclusions

Our current method of estimating self-discipline relies upon a self-reported survey administered once. There can be problems of both over- and under-reporting. We could take advantage of the continuous data students generate, and construct a more robust estimate of self-discipline. It may also be possible to consider self-discipline a latent construct, similar to what we do for knowledge in knowledge tracing, and simultaneously estimate both parameters. Broadening the stream of ITS information to include observable measures like homework submission, attendance, usage of tutor, opinion of teachers and parents, etc. would make this possible.

In conclusion, high self-discipline students have higher incoming knowledge, as substantiated from both KT model parameters and pre-test score. However, the impacts do not appear to be substantial, and tutor designers probably do not have to explicitly account for self-discipline. The higher self-discipline group makes better guesses and makes fewer slips, which implies that the higher self-discipline group is more careful and

detail oriented. The cumulative effect of learning, slip and guess makes the performance of higher self-discipline students better than that of their peers.

Acknowledgements

We would like to thank all of the people associated with creating the ASSISTment system listed at www.ASSISTment.org. We would also like to acknowledge funding from the National Science Foundation, the Fulbright Program for funding the second author and the US Department of Education and the Office of Naval Research for funding the other authors. All of the opinions expressed in this paper are those solely of the authors and not those of our funding organizations.

References

- [1] Angela L. Duckworth and Martin E.P. Seligman, *Self-Discipline Outdoes IQ in Predicting Academic Performance of Adolescents*, *Psychological Science*, 16, 939-944
- [2] Ivon Arroyo, Beverly Park Woolf, *Inferring learning and attitudes from a Bayesian Network of log file data*, 12th international conference on artificial intelligence in education, 2005.
- [3] Albert T. Corbett and John R. Anderson, *Knowledge tracing: Modeling the acquisition of procedural knowledge*, *User modeling and User adapted interaction* 4:253-278, 1995.
- [4] Joseph E. Beck, Kai-min Chang, Jack Mostow, Albert T. Corbett, *Does Help Help? Introducing the Bayesian Evaluation and Assessment Methodology*. *Intelligent Tutoring Systems 2008*: 383-394.
- [5] Baker, R., Walonoski, J., Heffernan, N., Roll, I., Corbett, A. & Koedinger, K. *Why students engage in "Gaming the System" behavior in interactive learning environments*. *Journal of Interactive Learning Research*. Chesapeake, VA: AACE. 19(2), 185-224
- [6] Angel de Vicente and Helen Pain, *Informing the Detection of the Students' Motivational State: An Empirical Study*, 6th International Conference, ITS 2002.
- [7] Joseph E. Beck and Kai-min Chang, *Identifiability: A Fundamental Problem of Student Modeling*, 11th International Conference, User Modeling, 2007.
- [8] Borghans, Lex, Meijers, Huub and Ter Weel, Bas, *The Role of Noncognitive Skills in Explaining Cognitive Test Scores* (November 2006). IZA Discussion Paper No. 2429. Available at SSRN: <http://ssrn.com/abstract=947088>
- [9] Tangney, J.,P., Baumeister, R.F., & Boone, A.L.(2004). *High self-control predicts good adjustment, less pathology, better grades, and interpersonal success*. *Journal of personality*, 72, 271-322.