Exploring Gritty Students' Behavior in an Intelligent Tutoring System

Erik Erickson[1], Ivon Arroyo[1], Beverly Woolf[2]

[1] Worcester Polytechnic Institute, Worcester MA 01609, USA
[2] University of Massachusetts Amherst, Amherst MA, 01003, USA
eerickson@wpi.edu

Abstract. This research focuses on determining whether a student's GRIT impacts their behavior within an intelligent tutoring system, towards developing better student models and feature sets that can help a tutor predict student behavior and determining whether computer tutors might foster improvements in students' grit, perseverance and recovery from failure. We use rare Association Rule Mining to explore how students' grit may be associated with students' behaviors within MathSpring, an intelligent tutoring system, as a first step.

Keywords: Grit, Perseverance, Student Models, Association Rule Mining.

1 Introduction

Studies have shown that grit is more predictive of life's outcomes compared to the "Big Five" personality model, which is a group of broad personality dimensions (e.g. conscientiousness, extraversion, agreeableness, and neuroticism [15]), but unlike IQ, the previous gold-standard predictor for life outcomes, grit may not be a static quality but one that can be developed [12]. Grit has become ubiquitous in the lexicon of public schools across America [20]. Educators are looking for answers to some lingering questions: "*Can students increase their grittiness?*" and "*How do students go about doing so?*". Gritty individuals can maintain high determination and motivation for a long time despite battling with 'failure and adversity'. Students can increase their grittiness through classroom activities [20]. Educators are interested in fostering growth in children, and would be interested in fostering grit in their students.

Our research focuses on how a student's grit and perseverance might impact behavioral patterns in a tutoring system, towards understanding how digital tutors might foster gritty-like behaviors, and in turn, grit assessments.

We move research on grit forward as a tool to refine student models in intelligent tutoring systems, by answering the following questions:

RQ#1. *Can we predict if a student is gritty or not by looking his/her behaviors? Here, grit is a target to predict, or a consequence.*

RQ#2. *Does the grit of a student influence student behavior inside a tutor? In which way(s)? Here grit is a cause or antecedent*

2 Method

Grit has typically been assessed using Duckworth's instrument of the Grit Scale [13], asking students to report on twelve Likert-scale questions. Some examples of questions are, "*I often set a goal but later choose to pursue a different one*" and "*Setbacks don't discourage me.*"

Our testbed is MathSpring, an intelligent tutoring system (ITS) that personalizes problems by assessing students' knowledge as well as effort and affect as they engage in mathematics practice online [5-7]. Students used MathSpring during class time over several days, as part of their regular mathematics class, and solved many math problems, while the system captured detailed event-level and problem-level information on their performance. These students also filled out a grit scale survey [8] that produced in an aggregate grit score.

2.1 Data Collection and Data Mining

Seventh grade students from two school districts participated in a research study. After combining the two datasets, there were 456 rows of Grit survey responses representing thirty-eight students. Sixty-eight students used MathSpring, producing 3,012 rows of data, each representing a student-math problem interaction. Variables were discretized into Booleans, indicating high/low or true/false. We created the negation of each variable (e.g., for GUESS, we also created a counterpart NoGUESS variable with the opposite truth value) to be considered also. Along with Guess, other variables included Hi/Low Grit, is/is not Solved, Hi/Low Mistakes, Hi/Low Hints, Yes/No Finished, Not/Likely Read (the problem).

We used Association Rule Mining to discover rules, a non-parametric method for exploratory data analysis, which finds associations that occur more frequently than expected from random sampling. The four critical parameters and minimum thresholds used are the following: Support 0.05, Confidence 0.84, Lift 1.15, Conviction 1.75. Last, we subjected the most important rules to a Chi-Square statistical test, those with solely "High Grit" or "Low Grit" as a consequent or antecedent.

3 Results

The mean Grit Score for the N=38 students in the sample was *M=3.07*, *SD=0.51, Median=3, Range= [1,5]*. This means the student grit assessment had some variability but the distribution is centered on a neutral grit value. A median split was done, classifying students as low or high grit, so that half of the students were considered gritty or not. Interestingly, we found that High-Grit students had much more activity, 71% of the student-problem interactions in the dataset vs. 29% for the non-gritty students. Table 4 shows the number and percent of cases for notable variable in detail, after the discretization process.

Due to a low support threshold of 0.05, thousands of rules were created. Only a selected subset of rules was chosen for interpretation, mainly those rules with a single consequent or antecedent, and those which met thresholds and had highest values for the metrics of confidence, conviction and lift.

Table 1. Name, number of Cases and Percent Cases for all Variables in the final dataset

Variable Name	N cases	% High (or True)	Counterpart Variable	N cases	% High (or True)
HiGrit	2146	71.25%	LowGrit	866	28.75%
GUESS	368	12.22%	NoGUESS	2644	87.78%
DNFINISH	261	8.67%	FINISHED	2751	91.33%
NOTREAD	86	2.86%	LIKELYREAD	2926	97.14%
isSolved	1655	54.95%	NotSolved	1357	45.05%
HiMistakes	1343	44.59%	LowMistakes	1669	55.41%
HiHints	822	27.29%	LowHints	2190	72.71%

A notable finding was that no rules with *LowGrit* as a consequent appeared at all according to our criteria specified in the parameter thresholds. This made us realize that, due to the much lower number of math problems seen by Low Grit students, the *confidence* for any rule with *LowGrit=1* as a consequent would be at chance level at 0.288 (as opposed to 0.5). We realized how the *confidence* metric is not very reliable in this case due to the imbalanced dataset. On the other hand, the metric that balances the rarity of the premises of a rule and their confidence is the '*conviction'* parameter. We thus set conviction as our first priority for selection of rules.

Table 5 shows the rules that had the highest conviction, confidence and lift. These rules also are the most complete rules (as generally subsequent rules that met the parameter thresholds had similar premises, but combined subsets of the propositions). **Rule A** is the rule with highest confidence, conviction and lift, and states that *if a student made a high amount of mistakes in a math problem, and asked for many hints as a way to help them solve the problem, then it means the student has a high level of Grit*. This joint condition happened in 19% of the total student-problem interactions examined. The significance of the effect for each rule was verified with a Chi-Square test by computing cross-tabulations between the premise being true/false vs. High/Low Grit ($p<0.0001$ for rules 1, 2, and 3).

Table 2. Grit as a Consequent: Association Rules with highest Conviction, Confidence, Lift

Rule	**Confidence**	**Conviction**	**Lift**	**Support**
Rule A. HiMistakes ^ HiHints → HiGrit *	0.89	2.56	1.25	0.19
Rule B. LowMistakes ^ isSolved → Low Grit *	0.45	1.29	1.56	0.10

* Significant difference at $p<0.0001$, χ^2 (1, N=3012)

On the other hand, no rules were found that met the thresholds of confidence, lift and conviction for *LowGrit* as a consequent. Still, we show the rule that has the best outcome for those metrics. The implication *LowMistakes ^ isSolved → Low Grit* has a confidence level of 0.45, which is low, however, it is higher than chance as stated earlier (chance level for any *LowGrit* row is 0.288). The rule suggests that if a student solves problems by making a low number of mistakes, then the student is NOT gritty.

Table 6 summarizes the found rules with Low/High Grit as a premise. This time, it was easier to find rules with *LowGrit* as an antecedent that met the thresholds of confidence, lift and conviction but not for *HiGrit*. Rule C is the main rule found for Low Grit as an antecedent (other similar rules are variations of this same effect), suggesting that if a student has low grit, then they will likely ask for few hints in a problem.

The rule that contains *HiGrit* as an antecedent is Rule D. While Rule D does not meet the lift and conviction thresholds we had set, it does meet the confidence threshold, and is the rule found with the highest values of confidence and conviction. This rule captures that *if a student is gritty, then the student will not quick-guess the correct answer to a problem.* Remember that guessing implies that a student entered many answers incorrectly and did not ask for help/hints, until they manage to solve it correctly (the multiple-choice format in most questions in MathSpring probably favors this type of disengagement behavior in general). We consider that students who guess are avoiding help when they should instead be asking for it, as they are answering incorrectly, as stated in previous research [1,2]. Rushing to get the right answer without fully understanding why, and avoiding seeking help.

Table 3. Grit as an Antecedent: Association Rules with highest Conviction, Confidence, Lift

Rule	Confidence	Conviction	Lift	Support
Rule C. Low Grit → Low Hints *	0.88	2.27	1.21	0.18
Rule D. Hi Grit → NoGUESS *	0.89	1.16	1.02	0.7

* Significant difference at $p<0.0001$, χ^2 (1, N=3012)

4 Discussion

This research starts unpacking how grit may be expressed in student behaviors inside an intelligent tutor, and on learning how fostering gritty-like behaviors might eventually improve a students' grit. In general, the results of Association Rule Mining suggest that there are differences students' behaviors depending on their assessed level of grit. Apparently, students who are gritty tend to neither quick-guess answers to problems, nor making lots of mistakes while avoiding help. At the same time, rules found with grit as a consequent suggest that if a student is in a situation of conflict, making mistakes but resolving them by asking for hints (or videos or examples), we can predict that the student has high grit. This is a desirable behavior when facing challenge in interactive learning environments, as specified by a review on help seeking and help provision in interactive learning environments [2].

It was harder to find Association Rules that associate students with low grit with behaviors (there are not as many systematic behavior patterns that could be associated to students of low grit). Still, the few rules found suggest that when a student has low levels of grit, they will seek for a low amount of hints. Conversely, the behavior that a student is NOT gritty is that he/she makes a low number of mistakes and eventually solves the problems correctly. Given the agency that MathSpring allows (more than most other learning environments) this does not necessarily mean that low-grit students tend to solve problems correctly (otherwise solve-on-first would have been part of the rules found). Students who skip problems or give-up will receive easier problems in an adaptive tutor. Also, students could choose material that is easier, or already mastered, to guarantee higher levels of success. Further analyses could help discern if this is the case, by analyzing the level of difficulty of the problems students received. Grit is a construct that will predetermine students to have different kinds of self-regulatory behaviors while learning in interactive learning environments.

Acknowledgements

This material is based upon work supported by the National Science Foundation under NSF award #1551594 INT: Collaborative Research: Detecting, Predicting and Remediating Student Affect and Grit Using Computer Vision. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

1. Aleven, V., McLaren, B., Roll, I., and Koedinger, K., Toward Meta-Cognitive Tutoring: A Model of Help-Seeking with a Cognitive Tutor. International Journal of Artificial Intelligence in Education (IJAIED), 2006. 16(2): 101- 128.
2. Aleven, V., Roll, I., McLaren, B., Ryu, E. J., and Koedinger, K. (2005) An Architecture to Combine Meta-Cognitive and Cognitive Tutoring: Pilot Testing the Help Tutor. Artificial Intelligence in Education: Supporting Learning Through Intelligent And Socially Informed Technology, 125: 17.
3. Aleven, V., Stahl, E., Schworm, S., Fischer, F., & Wallace, R.M. (2003). Help Seeking and Help Design in Interactive Learning Environments. Review of Educational Research, 73(2), 277-320
4. Arroyo, I. and Woolf, B. (2005) Inferring Learning and Attitudes from a Bayesian Network of Log File Data. In Proceedings of the 12th International Conference on Artificial Intelligence in Education, 33-40.
5. Arroyo, I., Beal, C., Murray, T., Walles, R., and Woolf, B. P. (2004) Web-Based Intelligent Multimedia Tutoring for High Stakes Achievement Tests Intelligent Tutoring Systems. In Proceedings of the 7th International Conference on Intelligent Tutoring Systems (ITS'04), Maceo, Brazil, 142-169.
6. Arroyo, I., Woolf, B. P., and Beal, C. R. (2006) Addressing Cognitive Differences and Gender During Problem Solving. International Journal of Technology, Instruction, Cognition and Learning, 4: 31-63.
7. Arroyo, I., Woolf, B. P., Burelson, W., Muldner, K., Rai, D., & Tai, M. (2014). A multimedia adaptive tutoring system for mathematics that addresses cognition, metacognition and affect. *International Journal of Artificial Intelligence in Education*, *24*(4), 387-426.
8. Corno, L., & Snow, R. E., Adapting teaching to individual differences among learners, in Handbook of research on teaching, M.C. Wittrock, Editor. 1986, MacMillan: New York.
9. Craig, S., Graesser, A., Sullins, J., and Gholson, B., Affect and learning: an exploratory look into the role of affect in learning with AutoTutor. Journal of Educational Media, 2004. 29(3): 241-250.
10. Csikszentmihalyi, M., Flow: the psychology of optimal experience. 1st ed. 1990, New York: Harper & Row. xii, 303 p.

11. D'Mello, S. and Graesser, A., Mind and body: Dialogue and posture for affect detection in learning environments, in Proceedings of the 2007 Conference on Artificial Intelligence in Education: Building Technology Rich Learning Contexts That Work. 2007, IOS Press. p. 161-168.
12. Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. Journal of Personality and Social Psychology, 92 (6), 1087-1101.
13. Duckworth, Angela & Quinn, Patrick. (2009). Development and validation of the Short Grit Scale (GRIT-S). Journal of personality assessment. 91. 166-74. 10.1080/00223890802634290.
14. Eskreis-Winkler, L., Gross, J. J., & Duckworth, A. L. (2016). Grit: Sustained self-regulation in the service of superordinate goals. *Handbook of self-regulation: Research, theory and applications. New York, NY: Guilford*.
15. Eskreis-Winkler, L., Shulman, E. P., Beal, S. A., & Duckworth, A. L. (2014). The grit effect: Predicting retention in the military, the workplace, school and marriage. *Frontiers in psychology*, *5*.
16. Goleman, D., Emotional Intelligence: why it can matter more than IQ. Bloomsbury. 1996, London.
17. Martin, N. and Halperin, S., Whatever it takes: How twelve communities are reconnecting out-of school youth. 2006, American Youth Policy Forum: Washington, D.C.
18. Pekrun, R., Goetz, T., Daniels, L., Stupinsky, R., and Perry, R., Boredom in achievement settings: Exploring control–value antecedents and performance outcomes of a neglected emotion. Journal of Educational Psychology, 2010. 102(3): 531-549.
19. Reed, J., Pritschet, B. L., & Cutton, D. M. (2013). Grit, conscientiousness, and the transtheoretical model of change for exercise behavior. *Journal of health psychology*, *18*(5), 612-619.
20. Saltman, K. J. (2014). The austerity school: Grit, character, and the privatization of public education. *symploke*, *22*(1), 41-57.
21. Schrum, L., & Levin, B. B. (2009). *Leading 21st-century schools: Harnessing technology for engagement and achievement*. Corwin Press.
22. Woolf, B., Arroyo, I., Muldner, K., Burleson, W., Cooper, D., Dolan, R., and Christopherson, R. (2010) The Effect of Motivational Learning Companions on Low Achieving Students and Students with Disabilities. In Proceeding of the10th International Conference on Intelligent Tutoring Systems (ITS'10), Pittsburgh, PA, 327-337.