

Relationships between math performance and human judgments of motivational constructs in an online math tutoring system.

Rurik Tywoniw¹[0000-0003-2885-5871], Scott A. Crossley¹[0000-0002-5148-0273],
Jaclyn Ocumpaugh², Shamyia Karumbaiah²[0000-0002-7920-4510], and Ryan Baker²

¹ Georgia State University, Atlanta GA 30303, USA
{rtywoniw1, scrossley}@gsu.edu

² University of Pennsylvania, Philadelphia PA, 19104 USA
penn.learninganalytics@gmail.com

Abstract. This paper explores how early grade school students' math performance relates to human ratings of students' affect, identity, and social awareness based on the content of messages to an online tutoring system avatar. There is an expanding body of research which investigates connections between these features and success in mathematics. This study used principle component analysis to identify four components related to motivational constructs. These components were examined using correlations with mathematics performance at three difficulty levels. Data from 572 students were examined, with results indicating little to no links between human judgments of motivational constructs and math performance. These findings have implications for how motivational constructs in math are evaluated and how they can predict mathematics performance.

Keywords: Intelligent Tutoring, principle component analysis, motivational constructs

1 Introduction

Recent research into math education has put emphasis on the effect of non-cognitive factors on math performance, such as math self-concept and motivation [1, 2, 3]. Intelligent tutoring systems provide an environment for self-paced learning and growth, and opportunities for interaction which contribute to development [4, 5]. In these online environments, positive sentiment towards the course is associated with positive course outcomes. Wen, Yang, & Rosé [6] examined sentiment analysis of postings in an online course, finding that latent affect features of positive impressions of the course were inversely proportional to course dropout rate. Slater et al. [4] found that students' self-perceptions of the value of math, their math self-concept, and interest in math each correlated with math performance. Crossley et al. [7] similarly found associations with math self-concept and math performance, also incorporating telemetric (click-stream) data from an online tutoring system as predictive of math identity. Missing from previ-

ous research is whether math performance is related to human judgments of math students' affect and identity. As students' use of online and intelligent interactive tutoring tools grows, it is useful to know if these constructs can be seen in student language and if these constructs are related to success. The current study thus asks the following questions: 1) Can human ratings of students' affect, identity, and social awareness be grouped into component macro-features related to motivational constructs? And 2) if such macro-features are discernible, do they relate to math performance?

2 Method

2.1 Data

Data were collected from *Reasoning Mind Foundations* by Imagine Learning, a blended learning platform for students in elementary grades. Students use this platform for self-paced engagement with math. Teachers use system data to monitor student performance and growth. Students can send emails to the Genie, a pedagogical agent who provides math help and encouragement. Messages sent to the Genie are responded to by employees of *Reasoning Mind* who maintain a consistent Genie persona. A more thorough description of the system is given in Khachatryan et al. [8]. The language sample for the analyses in this study come from the messages sent to the Genie tutor. These messages were aggregated into a single file for each student, allowing investigation of the content in individuals' messages, even when the average message by a given individual was short. Overall, the data in this study came from a sample of 572 elementary school students who used the *Reasoning Mind* platform between August 2016 and June 2017. Students attempted A-level (easiest), B-level (mid-level), and C-level (most difficult) math problems and wrote at least 50 words worth of combined messages to the Genie tutor. On average, students wrote 16 words per message. Students math performance scores are students' average performances on the A-, B-, and C-level problems. Data from this study are available upon request from the third and fifth author.

2.2 Human Ratings of Motivational Constructs

All aggregated message files were rated for evidence of students' motivational constructs in mathematics by two human raters. Students' messages were rated for fourteen different constructs, each on a scale from 1 to 5. These included affective features (Delight, Curiosity, Dejection, Engaged Concentration, Confusion, Frustration, Contempt), math identity features (Math Class Interest, Math Domain Interest, Math Self-Concept, And Non-Math Self-Concept), and social awareness features (Responsibility, Success, Cooperation). The two human raters were undergraduate students at a large university in the American South. The raters were trained and normed on similar tutoring messages from a previous data set. Their ratings were analyzed for intra-rater reliability using Multi-faceted Rasch Analysis [9]. Intra-rater reliability was satisfactory, with each rater exhibiting an infit of between .5 and 1.5 on each construct, indicating a satisfactory level of model fit and predictability without being invariant in their ratings.

2.3 Analysis

To answer the first research question, we performed dimensionality reduction using Principle Component Analysis (PCA), a statistical procedure which combines variables that are highly correlated into a smaller set of derived components. For inclusion into a component, a cut-off for the eigenvalues of $\lambda > .30$ was set, so only salient indices would be included in components. Each index was only included in the component in which it loaded highest. We calculated weighted component scores by multiplying each index by its respective eigenvalue in the component reported by the PCA. The results of the PCAs are discussed further in the Results section. To answer the second research question, the components resulting from the PCA were compared to math performance scores using Spearman's Rho correlations.

3 Results

3.1 Principal Component Analysis

The PCA was performed on the 13 variables from the raters' judgments of motivational constructs. The Kaiser–Meyer–Olkin test indicated that measuring of sampling adequacy (MSA) was sufficient at $MSA = .74$. Ten of the variables were retained in the analysis and reduced to four components with eigenvalues at or above 1.0. These four components accounted for 56.58% of the variance in human ratings of motivational constructs. These components were manually named based on indicator variables and are listed in Table 1. The component "Mood" relates to presence of features related to delight in math and the absence of features related to frustration with math. The component "Outcomes" related to the absence of a successful outlook regarding math and the presence of an outlook on math related to engaged concentration, cooperation, and confusion; all concepts related to success-in-the-making. The component "Attitude" relates to absence of contempt for math, and presence of interest in math class. Finally, the component "Declarativity" relates to general interest in the math domain and absence of curiosity.

3.2 Correlations between motivational constructs ratings and performance

Correlations between components of motivational constructs with math performance are presented in Table 2. Spearman's Rho was used as a test statistic because the data were not normally distributed. A conservative alpha value was set at .002 using Bonferroni Correction for multiple comparisons. Each of the three math performance scores at different difficulties were pairwise correlated with $\rho > .450$ ($p < .002$). Only two of the motivational components were pairwise correlated. Mood, which involved students' expression of either frustration or delight, correlated strongly with Attitude ($\rho = .478$, $p < .002$), which similarly involved students' expression of either contempt or interest in the math class. None of the components of motivational constructs were significantly correlated with math performance at any of the three levels.

Table 1. Components from the PCA on human judgments of motivational constructs

Component name and included variables	Percent of variance	Cumulative variance	Eigen loading for indices
1. Mood	25.25	25.25	
a. Frustration (-) (affect)			-.369
b. Delight (+) (affect)			.344
2. Outcomes	13.72	39.17	
a. Success (-) (social)			-.411
b. Engaged concentration (+) (affect)			.319
c. Confusion (+) (affect)			.477
d. Cooperation (+) (social)			.376
3. Attitudes	9.67	48.83	
a. Contempt (-) (affect)			-.457
b. Math Class Interest (+) (identity)			.354
4. Declarativity	7.74	56.58	
a. Math Domain Interest (+) (identity)			.706
b. Curiosity (-) (affect)			-.567

Table 2. Correlations between motivational construct components and math performance.

	A-level ⁺	B-level ⁺	C-level ⁺	Mood	Outcomes	Attitude
B-level ⁺	0.599*					
C-level ⁺	0.456*	0.487*				
Mood	-0.006	-0.004	-0.001			
Outcomes	0.001	-0.049	-0.004	-0.071		
Attitude	0.010	0.016	0.036	0.478*	0.027	
Declarativity	-0.038	-0.005	-0.111	-0.026	-0.052	0.075

* Significant at $p < .002$, ⁺ Level of difficulty of performance

4 Discussion

This paper described efforts to relate elementary level students' math performance to human ratings of affect, identity, and social awareness in their messages to a tutor. We successfully derived four components related to motivational constructs. Overall, there were no significant relationships between human judgments of motivational constructs in messages to an online tutoring avatar and math performance at three different level. This finding is in contrast with previous studies which have found more meaningful connections between motivational constructs and math performance [1, 3, 4, 5, 7]. However, this only implies that the effect on math performance of externally evaluated motivational constructs found in student writing may be mitigated by other factors which we could not measure, such as prior knowledge and tutoring environment features such as the content of tutor responses. These factors could be the subject of future studies. Considering the informal nature of the writing rated in this study, finer-grained metrics of affective- and identity-related features in language and telemetric data, to predict math achievement may also be effective.

Acknowledgements

This research was supported in part by NSF 1623730. Opinions, conclusions, or recommendations do not necessarily reflect the views of the NSF. We give special thanks to Lucile Petite and Rebecca Wood for their assistance in helping annotate the messages used in this study.

References

1. Spinath, B., Spinath, F., Harlaar, N., Plomin, R. (2006). Predicting school achievement from general cognitive ability, self-perceived ability, and intrinsic value. *Intelligence*, 34(4), pp. 363–374
2. Steinmayr, R., Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning & Individual Differences*, 19(1), pp. 80-90.
3. Arens, A. K., Marsh, H. W., Craven, R. G., Yeung, A. S., Randhawa, E., & Hasselhorn, M. (2016). Math self-concept in preschool children: Structure, achievement relations, and generalizability across gender. *Early Childhood Research Quarterly*, 36, 391-403.
4. Slater, S., Ocumpaugh, J., Baker, R., Lib, J., & Labrum, M. (2018). Identifying changes in math identity through adaptive learning systems use. In *Proceedings of the 26th International Conference on Computers in Education*.
5. Dowell, N. M. M., & Graesser, A. C. (2014). Modeling learners' cognitive, affective, and social processes through language and discourse. *Journal of Learning Analytics*, 1(3), 183-186.
6. Wen, M., Yang, D., & Rose, C. (2014, July). Sentiment Analysis in MOOC Discussion Forums: What does it tell us?. In *Educational data mining 2014*.
7. Crossley, S. A., Karumbaiah, S., Labrum, M., Ocumpaugh, J., & Baker, R. (2019). Predicting Math Success in an Online Tutoring System Using Language Data and Click-Stream Variables: A Longitudinal Analysis. *Proceedings of Language Data and Knowledge*, 70, (pp. 1-13). Open Access Series in Informatics (OASIs).
8. Khachatryan, G., Romashov, A., Khachatryan, A., Gaudino, S., Khachatryan, J., Guarian, K., & Yufa, N. (2014). Reasoning Mind Genie 2: An intelligent tutoring system as a vehicle for international transfer of instructional methods in mathematics. *International Journal of Artificial Intelligence in Education*, 24(3), 333-382.
9. Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions*, 16(2), 878.