JOURNAL OF LEARNING ANALYTICS

Volume 7(1), 19–37. http://dx.doi.org/10.18608/jla.2020.71.3

Predicting Math Identity Through Language and Click-Stream Patterns in a Blended Learning Mathematics Program for Elementary Students

Scott A. Crossley[1], Shamya Karumbaiah[2], Jaclyn Ocumpaugh[3], Matthew J. Labrum[4], Ryan S. Baker[5]

Abstract

This study builds on prior research by leveraging natural language processing (NLP), click-stream analyses, and survey data to predict students' mathematics success and math identity (namely, self-concept, interest, and value of mathematics). Specifically, we combine NLP tools designed to measure lexical sophistication, text cohesion, and sentiment with analyses of student click-stream data within an online mathematics tutoring system. We combine these data sources to predict elementary students' success within the system as well as components of their math identity as measured though a standardized survey. Data from 147 students was examined longitudinally over a year of study. The results indicated links between math success and non-cognitive measures of math identity. Additionally, the results indicate that math identity was strongly predicted by click-stream variables and the production of more lexically sophisticated and cohesive language. In addition, significant variance in math identity was explained by affective and cognitive variables. The results indicate that NLP and click-stream data can combine to provide insights into non-cognitive constructs such as math identity.

Notes for Practice

- **Overview**: Learning analytics research is beginning to focus on non-cognitive constructs that lend themselves to more immediate measures, including fine-grained measures of affective states and behavioural disengagement. More distal constructs have also been identified as important in the sociological and social-psychological research, including self-discipline (Duckworth & Selignman, 2005), social belonging (Walton & Cohen, 2011), and academic identities. Despite their known importance, there have been fewer publications that use student interactions (e.g., click-stream data and language production) to model these distal constructs.
- **Summary of contributions**: This study examines how self-concept, interest, and value relate to student behaviours within mathematics learning software. The study finds support for the notion that math identity constructs are correlated with math success and that language features in student texts are predictive of math identity constructs. The study finds some support for the use of click-stream variables to predict math identity, but these variables are generally less predictive. The study finds no evidence that math identity grows (or reduces) across the year of study analyzed here.
- **Key implications**: This study provides a foundation for examining non-cognitive variables using language features along with click-stream variables that can assist practitioners and software designers in identifying socially relevant math behaviours within online learning systems.

Keywords

Natural language processing, click-stream data, math success, math identity, longitudinal math development, corpus linguistics

Submitted: 05.04.2019 — **Accepted:** 17.12.2019— **Published:** 03.04.2020

Corresponding author [1]Email: sacrossley@gmail.com Address: Department of Applied Linguistics/ESL, Georgia State University, 25 Park Place, Suite 1500, Atlanta, GA 30303, USA

[2] Email: shamya@upenn.edu Address: Penn Center for Learning Analytics, University of Pennsylvania, 3700 Walnut St., Philadelphia, PA 19104

[3] Email: jlocumpaugh@gmail.com Address: Penn Center for Learning Analytics, University of Pennsylvania, 3700 Walnut St., Philadelphia, PA 19104

[4] Email: matthew.labrum@imaginelearning.com Address: Imagine Learning, 382 W. Park Circle, Suite 100, Provo, UT 84604

[5] Email: ryanshaunbaker@gmail.com Address: Graduate School of Education, University of Pennsylvania, 3700 Walnut St., Philadelphia, PA 19104

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

1. Introduction

As educational technology has become more complex, there has been increased research into the non-cognitive factors that shape and determine student choices. Over the last decade, researchers in learning analytics have paid close attention to non-cognitive constructs that lend themselves to more immediate measures, including fine-grained measures of affective states and behavioural disengagement, which have been modelled with click-stream data (i.e., interaction data from a learning system; Baker & Rossi, 2013; Baker & Ocumpaugh, 2014), physical and physiological sensor data (D'Mello & Kory, 2015), and language data (e.g., Dowell & Graesser, 2014). More distal constructs have also been identified as important in the sociological and social-psychological research, including self-discipline (Duckworth & Selignman, 2005), social belonging (Walton & Cohen, 2011), and academic identities, where a student values a particular academic subject as central to their self-concept (Osborne & Jones, 2011). Despite their known importance, there have been fewer publications that use student interactions (e.g., click-stream data and language production) to model these distal constructs. Thus, they remain a promising area of research for those interested in exploring the degree to which such constructs can be modelled and supported within educational technologies.

One reason that constructs related to identity have been less thoroughly researched may be related to methodological challenges. Many studies of identity constructs use qualitative methods that lend themselves better to small-scale analyses (e.g., discursive analyses of sociological interviews). These methods make it difficult to obtain standardized and large-scale ground-truth data for use in quantitative analysis. Other studies use lengthy surveys, which can also prove challenging for large-scale research. To date, few studies have been published that successfully infer coarse-grained survey measures from fine-grained log-file data, suggesting another potential issue: it may be difficult to model these mostly slowly shifting constructs without first parsing them into meaningful units of analysis (e.g., the discovery with models approach used in Ocumpaugh, San Pedro, Lai, Baker, & Borgen, 2016).

Still, there are dimensions to identity constructs that can make it more amenable to learning analytics approaches. For example, while some researchers do examine highly general constructs like learner identity (Baxter & Haycock, 2014; Solomon, 2007), academic identity (Reveles, Cordova, & Kelly, 2004; Nasir & Saxe, 2003), or orientation toward school (Osterman, 2000), others narrow the scope of the construct they are trying to capture to domain-specific identities, like math identity or physics identity (Syed, Azmitia, & Cooper, 2011). Likewise, while some research in math identity treats the construct very broadly (e.g., Cass, Hazari, Cribbs, Sadler, & Sonnert, 2011), other research on math identity decomposes the broad construct into finer-grained constructs such as student interest in mathematics (Renninger, 2009), mathematics self-efficacy/self-concept (Bong & Skaalvik, 2003; Pajares & Miller, 1994), and value of mathematics (Hitlin, 2003). These finer operationalizations make it possible to study math identity at scale using shorter surveys — providing ground truth that can then be used to develop finer-grained models.

In this study, we examine how constructs related to student math identity — the degree to which mathematics is considered an interesting and valuable component of a student's self-concept — might be modelled by using two different kinds of data involving their interactions within a mathematics software system. Specifically, we look at how their click-stream and language production data within a system can be used to model the math identity-related constructs of self-concept, interest, and value (operationalized using short, validated survey scales adapted from Ryan & Ryan, 2005). We also examine links between math identity constructs and success within the system (defined as percentage of beginning level math problems answered correctly) and examine changes longitudinally by examining pre-test and post-test data simultaneously. We do so within the context of Reasoning Mind, a blended-learning system used by over 100,000 U.S. elementary students each year as part of their regular mathematics instruction. In doing so, we investigate how student language in terms of text cohesion, structural patterns, and lexical sophistication, along with their system interactions, and time (pre-test and post-test) can predict math identity constructs. Our goal is to provide statistical models that can be used by practitioners and software designers to identify math behaviours within online learning systems.

1.1. Math Identity

Math identity — or the degree to which one considers oneself a "math person" — has become an area of interest among social scientists hoping to better understand what drives students to enter Science, Technology, Engineering, and Mathematics (STEM) fields (cf. Syed et al., 2011; Syed & Chemers, 2011; Watt et al., 2017). The literature on math identity has focused on various ways that students might frame self-definitions. For instance, students with high math identity have been described as perceiving themselves as capable of doing mathematics (i.e., self-concept; Hitlin, 2003; Renninger, 2009) or as deliberately engaging in mathematics learning (i.e., demonstrating interest in mathematics; Renninger, 2009). Math identity has also been described as a way of orienting oneself toward a specific set of values (Hitlin, 2003). In other words, math identity is a form of self-definition, where students see themselves as the type of person who is interested in mathematics and values their success in that topic.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

Modelling self-definitions may face challenges that more immediate constructs do not, particularly since they can also be susceptible to perceptual biases. For example, Bandura's (1977) research on self-efficacy discusses the role of self-attributional processes and suggests that a student's cognitive appraisal (self-evaluation of ability) may be susceptible to a form of confirmation bias where the student ignores demonstrable achievements and improvements when contrasted with a previously established self-definition (Bandura, 1977). In addition to issues related to bias, other challenges may make it difficult to model identity-related constructs. For example, researchers who have constructed self-concept measures have worked to ensure that this construct is deliberately separate from student performance (Gottfried, 1985; Steinmayr & Spinath, 2009).

In this study, we operationalize math identity as mathematics self-concept, mathematics interest, and mathematics value. We do so based on research from two different paradigms that suggest these constructs are related to identity and that, when broadly examined as non-cognitive constructs, can impact learning and motivation (Bandura, 1977; Cass et al., 2011; Eccles, 2009; Ryan & Ryan, 2005; Syed & Chemers, 2011). In addition to their appearance in Bandura's work, these constructs appear in Eccles's (2009) expectancy value theory, where self-efficacy (among a variety of other factors) is hypothesized to influence both intrinsic value (interest) and utility value (the usefulness of the task). Specifically, we measure self-concept, interest, and value using self-report scales adapted from Ryan and Ryan (2005), who examined relationships between these constructs and conditions likely to trigger stereotype threat effects in mathematics performance. We discuss each of our mathematics constructs briefly below.

1.2. Mathematics Self-Concept

Research in self-concept overlaps considerably with self-efficacy, and in some cases the term "identity" is also used to refer to a construct such as self-concept rather than the broader conceptualization we adopt above. All three constructs refer to the mental schema a person uses when calculating their ability to negotiate different challenges in their lives. In general, social psychologists are more likely to refer to the concept of identity when discussing issues related to social processes, while they are more likely to use the term self-concept when discussing internal mental processes (Schlenker & Weigold, 1989).

In education research, self-concept and self-efficacy are often used to discuss domain-specific evaluations (e.g., self-concept in mathematics), and they are sometimes used synonymously. However, some education researchers draw a careful distinction between these two constructs, limiting the term self-efficacy to self-evaluations of specific tasks, often specifying that it must be measured directly after the task has been completed (Bong & Skaalvik, 2003; Pajares & Miller, 1994). For example, they might measure self-efficacy by administering a Likert scale immediately after each mathematics problem, asking students to indicate their confidence in each answer. In this research tradition, self-concept is a broader measure of ability within the domain, where its meaning more closely approaches its use among social psychologists, who define it as a theory of self (e.g., Epstein, 1973).

Self-concept is believed to operate below the level of consciousness, guiding people's interpretations and expectations of external events (cf. Schlenker & Weigold, 1989), yet it may sometimes have contradictory effects. For example, in a situation where a student failed a task in a domain for which they have high self-concept, they might be more willing to attempt the task again than someone with low self-concept would be. Alternatively, they might interpret the task as flawed since their performance did not match the expectations created by their self-concept.

Both education researchers and social psychologists believe that self-concept develops from experience, so people with more shallow or limited experiences (novices) are likely to be more susceptible to changes in self-concept (Epstein, 1973). However, other research suggests that developmental issues may differentially influence younger children's self-concept in ways that contradict the development of self-concept in teenagers and adults. That is, constructs like self-concept often start high and show declines (rather than starting low and building with mastery experiences), and these declines sometimes occur as young as elementary school (Marsh et al., 1983) as student ability to self-evaluate matures (e.g., Stipek, 1981) and their self-concepts across different domains begins to diverge (e.g., math self-concept vs. verbal concept, as discussed in Marsh et al., 1985).

Academic self-concept tends to have a reciprocal positive relationship with achievement (Shavelson & Bolus, 1982). High self-concept can make students more likely to persist through difficult mathematics, leading to improved academic outcomes, while repeated failure could lower self-concept, particularly if a student did not have previous mastery experiences in mathematics to serve as a buffer.

1.3. Interest in Mathematics

Motivational research defines interest as the propensity to engage with a particular subject over time through both affective and cognitive components (Hidi & Renninger, 2006). Studies on the relationship of interest to self-concept have often found that self-concept drives intrinsic interest in a given subject (Bandura & Schunk, 1981; Sansone, Weir, Harpster, & Morgan, 1992), with theorists suggesting that as self-efficacy increases, students feel safer investing effort and attention in a particular

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

topic (Roberts & DelVecchio, 2000).

Researchers have identified several simple strategies that appear to increase interest in the classroom, such as creating more challenging tasks for students or adding variety to the ways in which a student is asked to perform a task. However, others caution that some of these strategies may only improve situational interest (e.g., Hidi & Renninger, 2006), suggesting that individual interest (intrinsic interest) is almost always self-driven, possibly because it seems to be fed by increased self-efficacy. Other researchers have found that interest is highly susceptible to contextual effects that vary from student to student (cf. Sansone et al., 1992). Researchers in career theory (e.g., Campbell & Hackett, 1986) have found that interest, like self-efficacy, is directly responsive to performance success and failure.

Interest is an important complement to self-concept when defining math identity, since its development is known to improve self-regulatory strategies (Hidi & Renninger, 2006) and other factors related to identity formation (Eccles, 2009). Students with stronger interest in a subject are more likely to persist when confronted with frustrating challenges (Fink, 1998; Hidi & Renninger, 2006; Prenzel, 1992), so that strengthening skills in mathematics is a self-feeding cycle. In particular, enjoyable or pleasant experiences with a subject are likely necessary to develop the persistence needed to become an expert in that subject (Eccles, 2009).

1.4. Value of Mathematics

Mathematics value is the degree to which a student thinks that mathematics is or will be useful to their life. Like self-concept and interest, value (utility) has been linked to motivation in several different research traditions. Domain-specific value (i.e., mathematics value) is thought to be highly influenced by prior knowledge, since limited knowledge can make it difficult for a student to connect the utility of a domain-specific task to more personal values (Renninger, 2009). Hitlin (2003) describes values as an enduring set of beliefs related to a person's goal structure, suggesting that people feel most authentic when their behaviours align with their values. Gecas (2000) extends this by introducing value-identities, which occur when a value becomes a core part of how a person defines themselves across different contexts.

Within the social psychology and education research, value has been shown to be influenced by self-concept (Chouinard, Karsenti, & Roy, 2007). In turn, much like interest, value positively influences the kind of goal-setting practices that increase effort and persistence (Chouinard et al., 2007). It also appears to substantially influence interest (Frenzel, Goetz, Pekrun, & Watt, 2010).

Several researchers have found that social influences can influence value (perhaps more than self-concept or interest). In particular, parents can have a substantial effect on a child's mathematics value (Chouinard et al., 2007; Frenzel et al., 2010; Harackiewicz, Rozek, Hulleman, & Hyde, 2012), and Renninger (2009) notes that, for younger learners, higher value of a domain is more likely when the values of teachers, parents, and peers align. Cumulatively, these findings suggest that value is often the last component of math identity to develop unless external influences (e.g., parents) are involved, but they also show that value can be more susceptible to social pressures or interventions.

1.5. Click-Stream Variables and Non-Cognitive Skills

Fine-grained click-stream data derived from student learning and learning contexts can be used to examine non-cognitive skills in learners (Beheshitha, Hatala, Gašević, & Joksimović, 2016; Pardo, Han, & Ellis, 2016). It has also become increasingly easy to collect in computer-based learning systems (Greene, Muis, & Pieschl, 2010), making it a mainstay of learning analytics (Gašević, Dawson, & Siemens, 2015). Click-stream data has led to increased opportunities to explore human learning behaviours in ways that were difficult to make explicit before. While click-stream data may sometimes provide an incomplete picture of student experiences with a learning system, it can provide real-time, fine-grained, unobtrusive access to student behaviours (Knight, Buckingham Shum, & Littleton, 2013).

Click-stream data has proven a useful addition to studies that rely exclusively on self-report for research on student behaviours. For example, Winne and Jamieson-Noel (2002) studied undergraduate student self-reports about study tactics and observed that learners tend to overestimate their use of study tactics. Dupeyrat & Marine (2005) also report that behavioural measures of effort are more predictive of learning than are student self-reports. Other studies have reported that observed outcomes as operationalized through click-stream data are more strongly associated with achievement than self-reported values because click-stream data has lower biases than subjective reflections (Zhou & Winne, 2012). Thus, many researchers are beginning to combine click-stream data and self-reported measures to model non-cognitive skills instead of relying on one or the other (Beheshitha et al., 2016; Pardo et al., 2016). In one study, McQuiggan, Mott, and Lester (2008) developed models of learner self-efficacy from university students' click-stream data, using self-reports as ground truth. In another example, Lust, Elen, and Clarebout (2013) collected self-reports of self-efficacy and achievement goal orientations from undergraduate students and associated these variables with learning strategies identified through click-stream data.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

While research related to these non-cognitive constructs are common in the context of mathematics education research, there have been relatively few efforts to model these constructs through click-stream data. Notably, these include models related to non-cognitive constructs in mathematics, including achievement goal orientation (Beheshitha et al., 2016) and self-regulation more generally (e.g., Pardo et al., 2016), while others have investigated math self-efficacy and interest (Lust et al., 2013). Many of these studies have focused on models related to fairly immediate measures. For instance, early work relied on a variety of sensor data to model a constellation of related non-cognitive constructs in mathematics including confidence, frustration, excitement, and interest (Cooper et al., 2009). Other work has tried to model less-immediate measures of such constructs. For instance, Ocumpaugh et al. (2016) used click-stream data produced by students in middle school to predict measures of high-school STEM interest and self-efficacy. Similarly, click-stream data during middle school has been used to predict students' choice of eventual college majors (San Pedro, Ocumpaugh, Baker, & Heffernan, 2014). More recent work highlights the crucial role of school demographics on the relationship between clickstream variables in an online math tutor and students' motivational measures indicating the influence of broader student social context while modelling non-cognitive factors (Karumbaiah, Ocumpaugh, & Baker, 2019).

1.6. Natural Language Processing (NLP) Variables and Non-cognitive Skills

NLP allows for the computational examination of the linguistic properties of text in order to understand natural language samples (e.g., student emails sent within an online tutoring system) and their relationships to practical items (e.g., math identity). Traditional NLP tools generally queried a text to better understand the level of complexity in terms of syntactic and lexical properties. Contemporary tools build on these features but also provide measurements of text organization and structure (e.g., cohesion) and sentiment analysis. NLP approaches have become more common in learning analytics studies in the last ten years, especially for adult learners, and are commonly used to better understand student writing (Allen, Likens, & McNamara, 2018; Botarleanu, Dascalu, Sirbu, Crossley, & Trausan-Matu 2018), text readability (Crossley, Skalicky, & Dascalu, 2019), and student success in massive open online classes (MOOCs; Crossley, Paquette, Dascalu, McNamara, & Baker, 2016; Wen, Yang, & Rosé, 2014a, 2014b) and other online learning environments (Crossley, Barnes, Lynch, & McNamara, 2017; Crossley, Karumbaiah, Labrum, Ocumpaugh, & Baker, 2019).

A few studies have examined how NLP features in student texts can be used to identify non-cognitive factors. However, much of this work focuses on adult learners. For instance, Moon, Potdar, and Martin (2014) used emotion terms and semantic similarity among participants to identify student leaders in two computer science MOOCs. Studies by Dascalu and colleagues examine how NLP features related to cohesion network analysis (CNA) can examine collaborative and participation patterns by students in online forums in MOOCs and other learning environments (Dascalu et al., 2018; Sirbu, Dascalu, Crossley, McNamara, & Trausan-Matu, 2019). While less common, at least one study has examined math identity using NLP features in elementary children (Crossley et al., 2018). In this study, language features in student writing were used to model student math identity in an online mathematics tutoring system. Crossley et al. examined relationships between student language and survey measures of math identity and found that NLP features were able to predict a significant amount of the variance in each math identity variable. Specifically, students who used more positive language were more likely to have a positive math identity while those who produced more negative language were more likely to have lower math identity. Students with stronger math identity also used more respectful language, less power-related language, and language that was calmer. Students with stronger math identity were also more likely to produce more sophisticated words. Crossley et al. also reported weak to medium relationships between mathematics success in the system and math identity.

1.7. Current Study

The current study builds on Crossley et al. (2018), which used NLP tools to model end-of-year math success (performance) within Reasoning Mind. Like the Crossley et al. study, we model three non-cognitive constructs related to math identity: math interest, math value, and math self-concept. However, we build on the original study in two ways. First, we include several click-stream variables as predictors of these math identity constructs, thus combining dispositional, log data, and text data. Second, the data in the current study is longitudinal (including both start-of-year and end-of-year language, math performance, and survey measures), allowing us to examine whether time is a predictor of math identity changes. By including these additional measures, we can examine differences across a year of study while also allowing us to better understand how individual variance, as determined by repeated participant data, can explain math identity beyond language and click-stream variables.

We use survey data to derive our math identity variables. To derive the language features for this study, we use NLP tools to extract language features from student emails to a virtual pedagogical agent within Reasoning Mind. As in the previous study, several NLP tools were applied, allowing us to extract language information related to text cohesion, lexical sophistication, and sentiment. Our click-stream data was extracted from the online tutoring data and focused on actions within

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

the system, entries into various modes of the system, and temporal data related to time spent in those modes. From the system, we also collected a single variable of math success by averaging students' success at answering math questions within the system. We collected data in two concurrent semesters (fall and spring) from students who typically used the system for several days each week, allowing us to track performance over time. Thus, in this study, we investigate three research questions:

RQ1. Are there relationships between math identity variables and mathematics success?

RQ2. Do math identity and mathematics success variables change over time (i.e., between pre- and post-year survey measures)?

RQ3. Are linguistic and click-stream features significant predictors of self-reported student traits related to mathematics value, interest, and self-concept over two semesters of study while controlling for student variance (i.e., pre- and post-year survey variance)?

2. METHOD

2.1. Reasoning Mind

We collected data from *Reasoning Mind Foundations*, a blended learning mathematics program for grades 2–5, distributed by Imagine Learning. *Foundations* students learn mathematics in an engaging, animated world at their own pace, while teachers use the system's real-time data to provide one-on-one and small-group interventions (Miller et al., 2015). The algorithms and pedagogical logic underlying Foundations (previously called *Genie 2*) are described in detail by Khachatryan et al. (2014), and a more thorough description of the system in general is given in Crossley et al. (2018).

Students using the *Foundations* curriculum start the system by entering the City Landscape (see Figure 1). From here, they click on virtual buildings, each of which correspond to a different mode of use (see Table 1). Not all modes in Foundations are instructional (e.g., the Mailbox and My Place modes), and some modes are further divided into smaller modules and content areas, which vary in their instructional capacity (e.g., the Guided Study and Game Room modes).

Figure 1. City landscape in Reasoning Mind's *Foundations*.

Students can send emails to the main character in the system, the Genie, a pedagogical agent who encourages students in their work throughout the system. The presence of the Genie and the ability to email it were implemented to increase both learning and enjoyment for students using the system (Khachatryan et al., 2014). Messages sent to the Genie are answered by Reasoning Mind employees who maintain a consistent persona by referencing an extensive biography of the Genie. The Genie is only able to reply to messages; emails to students are not initiated by the Genie.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

Table 1: Modes and Modules that Differentiate Student Experiences in *Foundations*

Mode	Module	Description
City Landscape		Main screen; the student selects a mode to enter from here and is taken here when exiting the mode.
Guided Study	Introduction Theory Notes Test Problems Review	The main study mode. Each objective begins with an Introduction and then the student works through the instructional content presented in the Theory module. A Notes Test is given to check the students' understanding of the theory before they proceed to the Problems modules and finally end with a Review.
Wall of Mastery		Additional practice mode; the students can select which content they want to practice.
Office		Assignments created by the teacher are accessed in this mode.
Homework		Homework assigned to the student is accessed and submitted in this mode.
The Genie		The Genie's house; this is an entertainment mode where students can see the Genie and play games.
Mailbox		Email functionality mode.
Shopping Mall		The place to purchase items, which can be used in the My Place mode.
Game Room	Math Race	A two-player game in which students race to the finish line by answering problems quickly and correctly.
	Division Game	A one-player game for practicing mental division calculation skills.
	Multiplication Game	A one-player game for practicing mental multiplication calculation skills.
	Dino Island	A one-player game for practicing proportions.
	Riddle Machine	A one-player game for practicing mathematics problems of slightly higher difficulty than would be encountered in Guided Study or Wall of Mastery.
Library		Reference material (e.g., Glossary, Genie's Rules)
My Place		An entertainment mode in which the student may interact with purchased items (e.g., decorate a room, watch short movies, read books).

2.2. Participants

The students sampled in this study came from a sample of 34,602 students who used Foundations from August 1, 2016 to June 17, 2017. The students were from 462 different schools located in 99 different districts, mostly in Texas. We included only those students who had attempted A-level math problems in both the fall and spring semester. We selected A-level problems because they are the lowest level math problems in the system and all students answer them, whereas B- and C-level problems are assigned less consistently. As an additional criterion for inclusion in our analyses, students included in this study needed to have written at least 50 words within the Genie email system (the minimum number of words needed to develop a linguistic profile). From the available data, 1,036 students met these criteria. However, we only included those who completed a math identity pre-test survey (given in the fall of 2016) and a math identity post-test survey (given in the spring of 2017) within the Genie email system. This subset of the data consisted of 147 students from 32 different schools in 11 districts in Texas.

2.3. Survey Data

The measures used in the present study consisted of three 4-point scales adapted from Mingle (2013). The same survey was administered at the start and end of the 2016–2017 school year. The first was mathematics self-concept, which comprised five items that captured the degree to which the student sees themself as a "math person" (e.g., "I have always been good at math").

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

The second was interest in math, which consisted of four items that captured intrinsic curiosity or enjoyment of math (e.g., "How much do you like math?"). The last scale measured the value of math and consisted of five items that captured the degree to which students find math to be useful (e.g., "How important is it to you to get good grades in math class?"). The survey scale was from 1 ("not at all important") to 4 ("very important"). Please see the Appendix for the math identity survey questions used in this study. The Cronbach α values of these scales were 0.72, 0.69, and 0.72 respectively. The survey responses under each construct were aggregated by averaging the individual responses. Examples of aggregated email chains sent by two 4th grade students (one classified as having low and one as having high math identity) are presented in Table 2.

Table 2. Example Emails for Reasoning Mind Students

Classified as low math identity	*Classified as high math identity*
I might dye my hair red or blue -Your bestie	Dear Genie, I love Reasoning Minds. How did you come up with it?
Hi Genie! What did you do for Valentine's Day? Your coolest friend	Hello, Genie! | Genie, I have been sending messages and I never got a reply. What happened? Are you okay? Did you get sick or something? If you tell me why you were not answering, I will forgive you.
Hi Genie! What are you gonna do for Saint Partick's day? Your Friend. Ps. you better wear green or i'll pinch you MWHAHAHAH	Genie, will you teach me how to do 2-digit by 2-digit, please? I keep on trying, but I just don't grasp the concept.
Dear Genie, I gave a bunch of candy to all my friends! Sadly, Its was raining. Your friend,	

2.4. Math Success

For this paper, math success is operationalized as students' average performance on A-level problems, the only problem level completed by all students using Reasoning Mind. B- and C-level problems are not just more difficult, but they are used in different ways by different teachers. Some teachers assign them to all students, other teachers either do not assign them at all or assign them only as challenge problems that students can optionally complete.

2.5. Genie Email Corpus

Our language sample for this analysis consisted of within-system email messages sent from the selected participants to the Genie. These messages were aggregated into a single file for each student, allowing us to investigate an individual student's linguistic activity more thoroughly, even in cases where the student sent a series of short messages. On average, students sent 8.73 (SD=10.2) messages to the Genie with an average length of 816.2 (SD=1546.4) characters.

Data cleaning procedures were implemented to reduce the amount of noise in the data. First, non-ASCII characters were removed, as these could interfere with the NLP tools. Second, all texts were automatically spell-checked and corrected using an open-source Python spelling correction library (autocorrect 0.3.0). After taking these preliminary data-cleaning steps, we removed random, non-mathematics symbols (e.g., "#," "@," and "&" but not currencies, percentages, timestamps, and ordinals). We also removed repeating words (e.g., "I love love love love love you Genie"), excessively long character strings (e.g., "agneggenepwbaegbepwxzmew"), and strings with repeating characters (e.g., "wooofffffrrrddd") and mixed-characters (e.g.,"$wofrd$"). Next, all non-dictionary, invalid words were removed from the data. We defined invalid words as those that could not be located within the synonym sets (synsets) in WordNet (Miller, 1995), and which also did not contain valid English phonotactic patterns. That is, a word was phonotactically invalid if it consisted of all consonants (always invalid), or if any pair of characters (digraph) in the word were invalid in the English language (e.g., "rlap" or "tsitl"). Finally, we removed all repeating, non-overlapping groups of words (e.g., "this word this word this word"). Only word groups of lengths two, three, and four were removed by this approach.

2.6. Natural Language Processing Tools

We selected NLP features that measure lexical sophistication, text cohesion, syntactic complexity sentiment analysis, and topic similarity as predictors of math success. We did so based on suggestions found in Crossley et al. (2018) and other studies that have examined links between NLP features and math success (Crossley et al., 2017, 2019). We used several NLP tools to calculate the selected linguistic features from our corpus of aggregated messages to the Genie, including the Tool for the

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

Automatic Analysis of Lexical Sophistication (TAALES; Kyle, Crossley, & Berger, 2018), the Tool for the Automatic Analysis of Cohesion (TAACO; Crossley, Kyle, & McNamara, 2016a), the Tool for the Automatic Analysis of Syntactic Sophistication and Complexity (TAASSC; Kyle & Crossley, 2017), and the SEntiment ANalysis and Cognition Engine (SEANCE; Crossley, Kyle, & McNamara, 2016b). In addition, we developed specific indices related to topics commonly discussed with the Genie email system using Latent Dirichlet Allocation (LDA). The features are discussed briefly below.

TAALES reports on several indices related to basic lexical information, lexical frequency, lexical range, lexical registers, word information features, and psycholinguistic variables including how long it takes to name a word aloud. For instance, the tool uses the Contemporary Corpus of American English (COCA; Davies, 2008) to compute the number of texts in which a word appears (i.e., range) to examine the specificity of words or the proportion of common bigrams in a text. The tool also reports on several phonological, orthographic, and phonographic neighbours that words must calculate how many near neighbours, based on sound or spelling, a word has. TAALES also reports on the strength of associations between words in the text to assess the expectancy of adjacent and nearby words.

TAACO incorporates a variety of classic and recently developed indices related to text cohesion to provide linguistic counts for both sentence and paragraph markers of cohesion. For example, TAACO calculates type token ratio (TTR) indices (which measure word repetition), sentence overlap indices that assess lexical and semantic similarity between sentences or paragraphs, and a variety of connective indices including additive connectives (i.e., and, moreover).

TAASSC measures large- and fine-grained clausal and phrasal indices of syntactic complexity and usage-based indices of syntactic sophistication. Syntactic sophistication indices focus on verb-argument constructions and their frequency and/or verb and argument association strength.

SEANCE is a sentiment analysis tool that contains several pre-developed word vectors that measure sentiment, cognition, and social order. SEANCE also provides a negation feature and a part of speech (POS) tagger. Examples of affective variables reports by SEANCE include positive and negative polarity metrics. Cognition examples include words related to attention (how much attention a word demands), terms of overstatement (i.e., words of emphasis), and words related to the mental processes associated with problem solving.

2.7. Click-stream Data and Grade-level Information

Reasoning Mind logs the interaction of the students in the system at the level of specific student actions within the interface. The log data includes a number of fields such as student identification (student ID, classroom ID, etc.), timestamp, mode, module type, event type (login, enter mode, problem shown, submit, etc.), and problem information (content type, dataset, try number, input type, student answer, result, etc.). For each student, we extracted features based on the actions in the 27 modes (e.g., City Landscape), the actions within the 11 module types of Guided Study mode (e.g., Introduction, Theory, Problems, Homework) and six content types (i.e., A-, B-, and C-level problems as well as Theory, Notes Test, and Speed Game). For each mode, we extracted three kinds of features: 1) the sum of entries to the mode (e.g., number of times the student entered the City Landscape or Guided Study); 2) the sum of actions performed within the mode/module/content type (e.g., number of actions performed in the Problems module type); 3) the total amount of time spent in each mode/content type (e.g., total time spent solving A-level problems). In addition, we calculated normalizing features (normalized against time in mode vs. time in system, etc.) and features measuring the number of hints, the number of virtual prizes purchased, and problem accuracies. In total, we mined 111 features from Reasoning Mind's log data, consisting of 110 click-stream features and student grade-level.

2.8. Statistical Analysis

Prior to analysis, all numeric scores were standardized. We used linear mixed effects (LME) models in R (R Core Team, 2018) using the *lme4* package (Bates, Maechler, Bolker, & Walker, 2015) to develop models of math identity scores over time (i.e., across the fall and spring semesters by using both the pre- and post-year survey scores). LME models were chosen because they offer statistical advantages over traditional repeated measures analyses of variance (RM ANOVAs). Specifically, LMEs account for both pooled and individual variance among students as opposed to just a pooled group mean; it does this by including students as random effects (i.e., assigning a unique intercept for each student, resulting in more accurate estimates based on individual student variation). These are reported as marginal R2 and conditional R2, which show the effects of fixed factors in explaining math identity constructs and the effects of variance from individual students on math identity constructs, respectively.

A unique LME model was generated for each math identity construct (i.e., three models). For each model, grade level, linguistics and affect features, and click-stream variables were treated as fixed effects (i.e., predictor variables), with 2nd grade

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

serving as the base-line for the grade-level feature. We also included students as random effects to measure the effects of successive sampling, which can affect the strength of relationships between dependent and independent variables.

To help control for over-fitting, we removed several variables prior to analysis. First, we conducted correlations (where possible) between the dependent variables and the independent variables. Any independent variable that did not demonstrate at least a small relationship with the dependent variable ($r \geq 0.100$) was removed from the analysis. Next, we checked for multicollinearity between the remaining independent variables using variance inflation factors (VIF) with a threshold set to 5 (i.e., high multicollinearity). All variables showing VIF above 5 were removed from the analysis and the remaining variables were used in the LME analysis. After variable pruning, we were left with 33 variables for modelling math self-concept, 32 for math interest, and 35 variables of math value. For each dependent variable, an initial LME model was run with all independent variables. After an initial model was constructed, we used a (backwards) stepwise variable selection technique to eliminate non-significant effects. The features remaining after stepwise variable selection process were used in the final models.

Several other statistical packages to aid in our construction and interpretation of our models. We used *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2017) to derive p-values from the models and to perform automatic backward elimination of variables in the LME models and the *MuMIn* package (Barton, 2018) to obtain two measures of variance explained: a marginal R2 measuring the variance explained by the fixed effects only, and a conditional R2 measuring the variance explained by the fixed and random effects combined.

3. RESULTS

3.1. Descriptive Statistics of Data

Descriptive statistics for the math success and identity variables are presented in Table 3. The descriptive statistics demonstrate that all variable means were above potential midpoints indicating that students scored relatively high on the A-level math problems and judged themselves to have higher than average math self-concept, interest, and value. For comparison, the mean A-level problem accuracy for a sample of ~11,000 students from the same year was 0.795, which is slightly lower than the mean score for the sub-sample in this analysis (0.838). Math value showed tendencies towards being non-normally distributed with skewness above 2.

Table 3. Descriptive Statistics for Math Success and Identity Variables

Variable	Mean	Standard Deviation	Skewness	Kurtosis
Math success (A-level problem accuracy)	0.838	0.105	–1.191	1.935
Math self-concept	2.946	0.704	–0.292	–0.722
Math interest	2.980	0.834	–0.614	–0.475
Math value	3.659	0.409	–1.865	4.588

3.2. Correlations between Math Identity Constructs and Mathematics Performance

To examine associations across the three math identity constructs and to assess links between math identity and mathematics scores, Pearson correlations were computed for both pre-test and post-test scores. For the purposes of this study, mathematics scores were limited to those derived from A-level mathematics problems within the Reasoning Mind system. The results (Table 4) indicate that all three math identity constructs (self-concept, interest, and value) were positively and significantly correlated with performance on A-level mathematics problems and with each other. Medium effects were reported for the relationship between mathematics performance and self-concept (0.308) while weak effects were reported for its relationship with interest (0.163) and value (0.181). Correlations between the math identity constructs themselves were also moderate (0.384–0.487), a result that aligns with their association in the literature.

Table 4. Correlations Between Mathematics Performance and Math Identity Variables

Variables	A-level accuracy	Self-concept	Interest
Self-concept	0.308**		
Interest	0.163*	0.432**	
Value	0.181*	0.384**	0.487**

* $p < .010$, ** $p < .001$

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

3.3. LME Model for Mathematics Self-Concept

An LME model predicting math self-concept as the dependent variable reported significant main effects for one click-stream variable and three linguistic features. The model reported a marginal R^2 of 0.062 and a conditional R^2 of 0.713, suggesting that most of the variance explained was in successive sampling of participants. Table 5 displays the estimates, percent of estimate, standard errors, *t*-values, and *p*-values for the fixed effects entered into this model.

Table 5. LME Model Predicting Math Self-Concept

Fixed effect	Variable type	Estimate	% of estimate	Std. Error	*t*	*p*
(intercept)		0.000		0.071	0.000	1.000
Time in A-level problems	CS	–0.189	0.398	0.047	–4.022	0.000
Additive connectives	TAACO	0.097	0.203	0.042	2.307	0.022
TTR (adjectives)	TAACO	–0.096	0.202	0.043	–2.244	0.026
Word range (COCA news) FW	TAALES	0.094	0.197	0.042	2.243	0.026

FW = function words; CS = click-stream

In general, the click-stream effects indicate that students who had higher self-concept spent less time solving A-level mathematics problems. Linguistically, students who had higher self-concept used more additive connectives and general function words (those found in a greater range of texts), results associated with more accessible text. These students also repeated more adjectives in their messages to the Genie, a pattern typical of narrative writing (Crossley, Kyle, & McNamara, 2016c).

3.4. LME Model for Mathematics Interest

An LME model predicting math interest as the dependent variable reported significant main effects for one click-stream variable and six linguistic features. The model reported a marginal R^2 of 0.116 and a conditional R^2 of 0.637, suggesting that most of the variance explained was in successive sampling of participants. Table 6 displays the estimates, percent of estimate, standard errors, *t*-values, and *p*-values for the fixed effects entered into this model.

Table 6. LME Model Predicting Math Interest

Fixed effect	Variable type	Estimate	% of estimate	Std. Error	*t*	*p*
(intercept)		0.000		0.069	0.000	1.000
Phonographic neighbours CW	TAALES	–0.140	0.156	0.050	–2.817	0.005
Problem solving words	SEANCE	–0.136	0.152	0.058	–2.366	0.019
Overstatement words	SEANCE	–0.131	0.147	0.054	–2.446	0.015
Proportion of bigrams (COCA magazine)	TAALES	0.125	0.140	0.054	2.305	0.022
Number of entries to My Place	CS	0.125	0.140	0.057	2.213	0.028
Positive terms (Hu Liu norms)	SEANCE	0.122	0.137	0.048	2.543	0.012
Word associations (delta p verb cue) SD	TAALES	0.114	0.128	0.047	2.451	0.015

CW = content words; SD = standard deviation; CS = click-stream

Only one click-stream feature was included in the interest model. Students who had higher interest entered their My Place rooms more often to store and access their purchased virtual prizes. Two sets of linguistic features were included in the interest model. Cognitively and affectively, students with higher interest scores used fewer problem-solving and overstatement terms coupled with more positive terms. This indicates that these learners used a greater number of positive words, fewer words of emphasis, and fewer words related to the mental processes associated with problem solving. In terms of linguistic structures, students with higher interest scores used words with fewer phonographic neighbours (an indication of greater lexical sophistication) and a greater proportion of expected bigrams (an indication of producing expected structural patterns). Students

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

with higher interest scores also showed higher standard deviations in the strength of associations between verbs and other words, which indicates that these students used phrases that demonstrated a greater variety of association strength (i.e., demonstrated greater phrasal knowledge).

3.5. LME Model for Mathematics Value

An LME model predicting math value as the dependent variable reported significant main effects for cognitive and linguistic features, but no click-stream variables were significant. The model reported a marginal R^2 of 0.085 and a conditional R^2 of 0.567, suggesting that most of the variance was explained in successive sampling of participants. Table 7 displays the estimates, percent of estimate, standard errors, *t*-values, and *p*-values for the fixed effects entered into this model.

Table 7. LME Model Predicting Math Value

Fixed effect	Variable type	Estimate	% of estimate	Std. Error	*t*	*p*
(intercept)		0.000		0.070	0.000	1.000
Semantic overlap between sentences (LSA)	TAACO	0.173	0.305	0.053	3.296	0.001
Word naming latencies standard deviations (FW)	TAALES	–0.134	0.236	0.050	–2.694	0.008
Attention words	SEANCE	0.145	0.255	0.050	2.906	0.004
Construction frequency types (COCA fiction)	TAASSC	0.116	0.204	0.048	2.417	0.016

FW = function words; LSA = latent semantic analysis

Linguistically, students who had higher scores for mathematics value produced text that was more semantically cohesive and produced more frequent syntactic constructions (i.e., more frequent verb argument constructions) indicating greater coherence between sentences and more frequent constructions that should make a message easier to read. The function words they produced also showed lower variance in word-naming scores, a timing measure that indicates that the students were producing fewer words at both the easy and more sophisticated ends of the extreme. In terms of cognitive variables, students with higher mathematics value scores used more attention words, a category operationalized as how much cognitive attention a word demands.

4. DISCUSSION

The purpose of this study was to examine links between language and click-stream variables in an educational software system for mathematics and three components of math identity, as measured by student survey responses. In addition, the study examined links between math identity variables and success within the system. At least one previous study had examined math identity using language features using a cross-sectional data analyses (Crossley et al., 2018) and several others have examined the relationship between click stream data and math identity (Slater, Ocumpaugh, Baker, Lib, & Labrum, 2018; Karumbaiah, Ocumpaugh, & Baker, 2019; Karumbaiah, Ocumpaugh, Labrum, & Baker, 2019). The current study builds on these studies modelling these constructs using a combination of click-stream data and language data and looking at changes in math identity across the academic school year.

Overall, we find similarities across math identity constructs as reported by correlations among the variables. We also find small to medium correlations among the math identity variables and mathematics success, with self-concept showing the strongest overlap with within-system mathematics problem solving. Linguistically, all math identity variables were at least partially predicted by lexical sophistication features (e.g., those identified by TAALES). Additionally, cohesion features (e.g., those identified by TAACO) were significant predictors in models of value and self-concept while language features related to sentiment and/or cognition (e.g., those identified by SEANCE) were significant predictors in models of value and interest.

The patterns reported in our linear mixed effects (LME) models provide a language profile for students with higher self-reported math identity. Specifically, these students produce more sophisticated words (i.e., words with fewer phonographic neighbours), but the range of sophistication for their function words was lower (i.e., words with lower variance in word naming speeds). They also produce text that is more cohesive, as demonstrated by a higher rate of overlap of semantic ideas, the use of more connectives, and a greater repetition of words. These students also followed expected structural patterns (e.g., patterns found in large-scale language corpora produced by proficient adult writers) including the production of less specific function

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

words, more expected bigrams, and more frequent verb argument constructions. In terms of sentiment and cognition variables, students with greater math identity used more positive terms, fewer words related to problem solving, and fewer words indicating emphasis and cognitive attention.

Fewer click-stream features were included in the models and, in some cases (math value), no click-stream features were significant predictors of math identity constructs. It might seem counter-intuitive at first that these features indicated that more successful students were spending less time working A-level problems (for self-concept) and more time in their virtual rooms (interest). However, less time in A-level problems means that students completed these more basic questions more quickly and were able to spend more time in other areas of the system, including their virtual rooms.

One limitation of this work is that the associations found in the LME models are relatively weak. As the marginal R2 values show, the language and click-stream features explained between 6% and 12% of variance in the math identity constructs. A larger portion of the variance (48% and 68%) was explained by random variance associated with the participants, much of which is likely the result of individual differences among the students.

Additionally, as reported in the LME models, time (fall pre-test and spring post-test) was not predictive of self-reported math identity variables, indicating no substantial links between math identity and time among the students in our sample (i.e., no longitudinal relationship) when time was co-varied with click-stream and linguistic features. To examine if differences existed between the math identity variables as a function of time, post-hoc paired t-tests were conducted for each variable (self-concept, interest, and value) and are reported here. The post-hoc analyses found no significant differences ($p > 0.05$) as a function of time for any of the variables. These results are encouraging, since it may be normal to see decreases in self-concept during this developmental phase (Marsh et al., 1983; Marsh et al., 1985; Stipek, 1981). Instead, these students seem to be maintaining relatively high values, which might suggest that Reasoning Mind is providing students with effective metacognitive support.

In the absence of time effect, our findings provide a greater understanding of how language and click-stream variables interact with math identity constructs and provide a pathway for future analyses. We discuss each model in turn below as well as discuss potential explanations for the variance explained by the participants themselves.

4.1. Mathematics Self-Concept

The first model we constructed was for self-concept, defined as a broad measure of confidence in one's mathematics ability related to one's theory of self. Our click-stream and language features explained 6% of the variance, indicating a weak correlation with self-concept while participant repetition explained over 60% of the variance (i.e., the successive measurements of participants was strongly correlated with their survey scores).

The strongest predictor in the LME model was time spent in A-level content; however, the coefficient was negative, indicating that students who spent less time in A-level content had higher self-concept. This indicates that students with greater self-reported ability spent less time on lower-level problems. This is also supported in the correlation analysis, which demonstrated medium relationships between self-concept scores and success in A-level content.

In terms of language use, students with higher self-reported self-concept produced text that had more cohesive features (i.e., additive connectives), tended to repeat adjectives, and used function words that were less specific (i.e., commonly used across texts). These linguistic features provide a profile of higher self-concept students as those who produce language that is more cohesive and easier to process. This may indicate that these students with higher self-concept (i.e., theory of self) produce emails that are more accessible/readable.

4.2. Mathematics Interest

Our second model explored math interest, which can be defined as the propensity to engage with a particular subject over time, with both affective and cognitive components. Like our self-concept model, the greatest amount of variance (62%) was explained by the random factor of participant (i.e., successive sampling). The fixed factors in the model explained around 12% of the variance in the math interest scores, indicating a medium effect size.

From a lexical perspective, students with higher interest used more sophisticated words, as demonstrated by their use of words with fewer phonographic neighbours and words with a greater range of associations (indicating knowledge of more potential associations). At the same time, these students produced more bi-grams that are commonly found in language, which suggests that they had knowledge of a greater variety of phases than students who self-reported lower math interest scores. Thus, linguistically, these students produced more sophisticated words within expected lexical frameworks.

Sentiment and cognition variables were also significant predictors of math interest. Specifically, students with higher math interest used more positive terms and fewer problem-solving and overstatement terms. These findings indicate that students with higher math interest were more positive in general when interacting with Genie and likely did not ask the Genie for

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

problem solving help or overstate issues by using emphatic words. Thus, these students likely sought less content help from the Genie.

Lastly, students with higher math interest spent more time in My Place, a space dedicated to leisure and entertainment, where students can furnish and redecorate a room, read books, or watch movies. We interpret this as indicating that students with higher math interest were more likely to have earned the points required to indulge in the activities related to this virtual space.

4.3. Math Value

Our third model examined links between language and click-stream variables and math value, which can be defined as the degree to which a student thinks that mathematics is or will be useful to their life. The model for math value explained about 9% of the variance based on language features alone (i.e., no click-stream variables were included), a weak to medium effect size.

The strongest predictor of math value was semantic overlap between sentences, indicating that students with higher math value produced more cohesive text. These students also used more words that led to cognitive attention, demonstrating an implicit use of words that garner interest. These students also showed less variety in their production of easy and difficult words (in terms of naming response times) indicating the ability to produce both less and more sophisticated words. Lastly, these students used more frequent verb argument construction types (i.e., more common syntactic structure). In general, this provides a language profile of students with higher math value as those that produce messages to the Genie that are more cohesive and easier to process (i.e., more common syntactic constructions and less variety in word difficulty).

4.4. General Trends

Overall, the three models begin to converge on a student with higher math interest as a student who writes more cohesively (i.e., greater overlap of ideas and greater use of cohesive features) and tends to write in expected patterns (i.e., more expected bi-grams, more general terms, and the use of more frequent verb-argument constructions). In addition, these students use more sophisticated words, but not necessarily a greater variety of sophisticated words. This finding indicates that these students have begun to develop patterns of text cohesion that are common in elementary students (Bereiter & Scardamalia, 1987; Berninger, Fuller, & Whitaker 1996) as well as beginning to norm their writing patterns to that of more proficient English writers. Thus, these students may be converging on the probabilistic language trends found in proficient writers, while students with lower math identity may not have acquired or noticed these patterns. Noticing these patterns may be related to general abilities to process information and extract information and this general ability may inform mathematics ability as well, predicting the significant correlations we see between mathematics success and math identity variables. More and different data, such as general tests of cognition, however, is needed to test these assumptions.

4.5. Practice and Policy Implications

The models constructed here provide practitioners and software designers with guideposts to begin identifying behaviours (linguistic and otherwise) within online learning systems that may indicate whether a student has high or low math identity. Combined with work previously presented on these issues (Crossley et al., 2017, 2018, 2019; Karumbaiah, Ocumpaugh, & Baker, 2019; Karumbaiah, Ocumpaugh, Labrum, & Baker, 2019; Slater et al., 2018), these models identify a range of features that might help to identify students who are most in need of support. Currently, the findings suggest that low-level linguistic patterns (e.g., those associated with the use of function words, connectives, TTR, lexical sophistication, and affect) and high-level linguistic patterns (semantic similarity and verb-argument constructions) may be stronger predictors of math identity than most click-stream data, which might be particularly useful information for learning systems that are unable to get students to complete surveys of math identity, but still wish to identify those students who are in most need of support.

The weaker performance of the click-stream variables is unexpected, but given the contrary effects that self-concept might have on behaviours and the developmental changes in identity that students are learning to manage, these weak associations are not entirely surprising. That is, if students who suffer from low math identity respond differently in their interactions with the system, these kinds of behaviours will be more difficult to classify cleanly. Practitioners and designers of learning systems should be aware of this result if they are hoping to support the development of stronger math identity among their students.

More research is needed to improve these models, and one obvious way forward is to revisit previous research on the demographic effects on the construction of math identity. In addition to the early work on how self-definitions are known to vary by gender, as in Bem's (1974) research, recent work shows that school-level differences in demographics may influence whether the students with high or low math identity are most likely to make use of system help-features (Karumbaiah, Ocumpaugh, & Baker, 2019). As language, like help-seeking, is culturally influenced, such explorations could help us to improve these models.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

5. CONCLUSION

This study examines how three constructs related to math identity — self-concept, interest, and value — are related to student behaviours within mathematics learning software. We find support for the notion that math identity constructs are correlated with math success as well as support that language features in student texts are predictive of math identity constructs. We find some support for the use of click-stream variables to predict math identity, but these variables are generally less predictive. Interestingly, we find no evidence that math identity grows (or reduces) across the year of study analyzed here. It may be that changes in math identity are incremental and longer timeframes are necessary to see changes. Alternatively, it could be that the experience of using this learning system weakens the drop in math identity often seen in young learners.

This study provides a foundation for examining non-cognitive variables using language features along with click-stream variables that can assist practitioners and software designers in identifying socially relevant math behaviours within online learning systems. While the current analysis explained only a small amount of the variance, we take solace in the notion that the variance explained was significant and meaningful in terms of effect sizes. We presume that with different metrics of math interest that are not self-reported, more advanced language features, and the inclusion of individual differences on the part of participants, we may be able to explain a greater amount of variance in future studies.

Acknowledgments

We would like to thank Tess Carithers, Leigh Mingle, Stefan Slater, and Victor Kostyuk for early work on this project that helped lead to this study. We would also like to thank the many staff members of Reasoning Mind who helped with data collection and to our anonymous reviewers for their very thoughtful assistance. Any mistakes are, of course, our own.

Declaration of Conflicting Interest

The data was collected in research partnership with Reasoning Mind, with whom Matthew J. Labrum was employed by at the time (and is now employed by Imagine Learning). The company's assets have since been bought by Imagine Learning and the Reasoning Mind Foundations program is no longer available as a product.

Funding

This research was supported in part by NSF grant number 1623730. Opinions, conclusions, or recommendations do not necessarily reflect the views of the NSF.

References

Allen, L. K., Likens, A. D., & McNamara, D. S. (2018). A multi-dimensional analysis of writing flexibility in an automated writing evaluation system. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (LAK '18), 5–9 March 2018, Sydney, NSW, Australia (pp. 380–388). New York: ACM. https://doi.org/10.1145/3170358.3170404

Baker, R. S., & Rossi, L. M. (2013). Assessing the disengaged behavior of learners. In R. Sottilare, A. Graesser, X. Hu, & H. Holden (Eds.), *Design recommendations for intelligent tutoring systems* (pp. 155–166). Orlando, FL: U.S. Army Research Lab.

Baker, R., & Ocumpaugh, J. (2014). Interaction-based affect detection in educational software. In R. A. Calvo, S. K. D'Mello, J. Gratch, & A. Kappas (Eds.), *The Oxford handbook of affective computing*. Oxford, UK: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199942237.013.009

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215. https://doi.org/10.1037/0033-295x.84.2.191

Bandura, A., & Schunk, D. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology, 41*(3), 586–598. https://doi.org/10.1037/0022-3514.41.3.586

Barton, K. (2018). *MuMIn: Multi-Model Inference*. R package version 1.42.1.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1–9.

Baxter, J. A., & Haycock, J. (2014). Roles and student identities in online large course forums: Implications for practice. *The International Review of Research in Open and Distributed Learning, 15*(1). https://doi.org/10.19173/irrodl.v15i1.1593

Beheshitha, S. S., Hatala, M., Gašević, D., & Joksimović, S. (2016). The role of achievement goal orientations when studying effect of learning analytics visualizations. *Proceedings of the 6th International Conference on Learning Analytics and Knowledge* (LAK '16), 25–29 April 2016, Edinburgh, UK (pp. 54–63). New York: ACM.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

https://doi.org/10.1145/2883851.2883904

Bem, S. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42*(2), 155–162.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written communication*. Hillsdale, NJ: Lawrence Erlbaum.

Berninger, V., Fuller, F., & Whitaker, D. (1996). A process approach to writing development across the life span. *Educational Psychology Review, 8,* 193–218. https://doi.org/10.1007/bf01464073

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review, 15*(1), 1–40.

Botarleanu, R., Dascalu, M., Sirbu, D., Crossley, S. A., & Trausan-Matu, S. (2018). ReadME: Generating personalized feedback for essay writing using the ReaderBench framework. In H. Knoche, E. Popescu, & A. Cartelli (Eds.), *The Interplay of Data, Technology, Place and People for Smart Learning: Proceedings of the 3rd International Conference on Smart Learning Ecosystems and Regional Development* (pp. 133–145). Springer. https://doi.org/10.1007/978-3-319-92022-1_12

Campbell, N., & Hackett, G. (1986). The effects of mathematics task performance on math self-efficacy and task interest. *Journal of Vocational Behavior, 28*(2), 149–162. https://doi.org/10.1016/0001-8791(86)90048-5

Cass, C. A., Hazari, Z., Cribbs, J., Sadler, P. M., & Sonnert, G. (2011). Examining the impact of mathematics identity on the choice of engineering careers for male and female students. *2011 Frontiers in Education Conference (FIE)*, Rapid City, SD, 2011 (pp. F2H-1–F2H-5). IEEE Computer Society. https://doi.org/10.1109/fie.2011.6142881

Chouinard, R., Karsenti, T., & Roy, N. (2007). Relations among competence beliefs, utility value, achievement goals, and effort in mathematics. *British Journal of Educational Psychology, 77*(3), 501–517. https://doi.org/10.1348/000709906x133589

Cooper, D. G., Arroyo, I., Woolf, B. P., Muldner, K., Burleson, W., & Christopherson, R. (2009). Sensors model student self-concept in the classroom. In G.-J. Houben, G. McCalla, F. Pianesi, & M. Zancanaro (Eds.), *International Conference on User Modeling, Adaptation, and Personalization* (UMAP 2009), 22–26 June 2009, Trento, Italy (pp. 30–41). Springer. https://doi.org/10.1007/978-3-642-02247-0_6

Crossley, S. A., Barnes, T., Lynch, C., & McNamara, D. S. (2017). Linking language to math success in a blended course. In X. Hu, T. Barnes, A. Hershkovitz, & L. Paquette (Eds.), *Proceedings of the 10th International Conference on Educational Data Mining* (EDM2017), 25–28 June 2017, Wuhan, China (pp. 180–185). International Educational Data Mining Society.

Crossley, S. A., Kyle, K., & McNamara, D. S. (2016a). The tool for the automatic analysis of text cohesion (TAACO): Automatic assessment of local, global, and text cohesion. *Behavior Research Methods, 28*(4), 1227–1237. https://doi.org/10.3758/s13428-015-0651-7

Crossley, S. A., Kyle, K., & McNamara, D. S. (2016b). Sentiment analysis and social cognition engine (SEANCE): An automatic tool for sentiment, social cognition, and social order analysis. *Behavior Research Methods, 49*(3), 803–821. https://doi.org/10.3758/s13428-016-0743-z

Crossley, S. A., Kyle, K., & McNamara, D. S. (2016c). The development and use of cohesive devices in L2 writing and their relations to judgments of essay quality. *The Journal of Second Language Writing, 32,* 1–16.

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. In M. Eskevich, G. de Melo, C. Fäth, J. P. McCrae, P. Buitelaar, C. Chiarcos, B. Klimek, & M. Dojchinovski (Eds.), *Proceedings of the 2nd Conference on Language Data and Knowledge* (LDK 2019)*,* 20–23 May 2019, Leipzig, Germany (pp. 1–13). Open Access Series in Informatics (OASICS), vol. 70. Retrieved from https://drops.dagstuhl.de/opus/portals/oasics/index.php?semnr=16105

Crossley, S., Ocumpaugh, J., Labrum, M., Bradfield, F., Dascalu, M., & Baker, R. S. (2018). Modeling math identity and math success through sentiment analysis and linguistic features. In K. E. Boyer & M. Yudelson (Eds.), *Proceedings of the 11th International Conference on Educational Data Mining* (EDM2018), 16–20 July 2018, Buffalo, New York, USA (pp. 11–20). *International Educational Data Mining Society*.

Crossley, S. A., Paquette, L., Dascalu, M., McNamara, D., & Baker, R. (2016). Combining click-stream data with NLP tools to better understand MOOC completion. In D. Gašević & G. Lynch (Eds.), *Proceedings of the 6th International Conference on Learning Analytics and Knowledge* (LAK '16), 25–29 April 2016, Edinburgh, UK (pp. 6–14). New York: ACM. doi:10.1145/2883851.2883931

Crossley, S. A., Skalicky, S., & Dascalu, M. (2019). Moving beyond classic readability formulas: New methods and new models. *Journal of Research in Reading, 42*(3–4), 541–561. https://doi.org/10.1111/1467-9817.12283

D'Mello, S. K., & Kory, J. (2015). A review and meta-analysis of multimodal affect detection systems. *ACM Computing*

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

Surveys (CSUR), 47(3), https://doi.org/10.1145/2682899

Davies, M. (2008). The Corpus of Contemporary American: 450 million words, 1990–present. Retrieved from http://corpus.byu.edu/coca/

Dascalu, M., Sirbu, M. D., Gutu-Robu, G., Ruseti, S., Crossley, S. A., & Trausan-Matu, S. (2018). Cohesion-centered analysis of sociograms for online communities and courses using *ReaderBench*. In V. Pammer-Schindler, M. Pérez-Sanagustín, H. Drachsler, R. Elferink, & M. Scheffel (Eds.), *Lifelong Technology-Enhanced Learning* (EC-TEL 2018). Lecture Notes in Computer Science, vol. 11082. Springer, Cham. https://doi.org/10.1007/978-3-319-98572-5_59

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. https://doi.org/10.18608/jla.2014.13.18

Duckworth, A. L., & Seligman, M. E. P. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science, 16*(12), 939–944. https://doi.org/10.1111/j.1467-9280.2005.01641.x

Dupeyrat, C., & Mariné, C. (2005). Implicit theories of intelligence, goal orientation, cognitive engagement, and achievement: A test of Dweck's model with returning to school adults. *Contemporary Educational Psychology, 30*(1), 43–59. https://doi.org/10.1016/j.cedpsych.2004.01.007

Eccles, J. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educational Psychologist, 44*(2), 78–89. https://doi.org/10.1080/00461520902832368

Eccles, J. (2011). Gendered educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *International Journal of Behavioral Development*, *35*(3), 195–201. https://doi.org/10.1177/0165025411398185

Epstein, S. (1973). The self-concept revisited: Or a theory of a theory. *American Psychologist, 28*(5), 404–416. https://doi.org/10.1037/h0034679

Fink, R. P. (1998). Interest, gender, and literacy development in successful dyslexics. In L. Hoffmann, A. Krapp, K. A. Renninger, & J. Baumert (Eds.), *Interest and Learning: Proceedings of the Seeon Conference on Interest and Gender* (pp. 402–407). Kiel, Germany: IPN.

Frenzel, A. C., Goetz, T., Pekrun, R., & Watt, H. M. (2010). Development of mathematics interest in adolescence: Influences of gender, family, and school context. *Journal of Research on Adolescence*, *20*(2), 507–537. https://doi.org/10.1111/j.1532-7795.2010.00645.x

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends, 59*(1), 64–71. https://doi.org/10.1007/s11528-014-0822-x

Gecas, V. (2000). Value identities, self-motives, and social movements. *Self, Identity, and Social Movements*, *13*, 93–109.

Gottfried, A. E. (1985). Academic intrinsic motivation in elementary and junior high school students. *Journal of Educational Psychology, 77*(6), 631–645. https://doi.org/10.1037/0022-0663.77.6.631

Greene, J. A., Muis, K. R., & Pieschl, S. (2010). The role of epistemic beliefs in students' self-regulated learning with computer-based learning environments: Conceptual and methodological issues. *Educational Psychologist, 45*(4), 245–257. https://doi.org/10.1080/00461520.2010.515932

Harackiewicz, J., Rozek, C., Hulleman, C., & Hyde, J. (2012). Helping parents to motivate adolescents in mathematics and science: An experimental test of a utility-value intervention. *Psychological Science, 23*(8), 899–906. https://doi.org/10.1177/0956797611435530

Hidi, S., & Renninger, K. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4

Hitlin, S. (2003). Values as the core of personal identity: Drawing links between two theories of self. *Social Psychology Quarterly*, *66*(2), 118–137. https://doi.org/10.2307/1519843

Karumbaiah, S., Ocumpaugh, J., & Baker, R. S. (2019). The influence of school demographics on the relationship between students' help-seeking behavior and performance and motivational measures. In C. F. Lynch, A. Merceron, M. Desmarais, & R. Nkambou (Eds.), *Proceedings of the 12th International Conference on Educational Data Mining* (EDM2019), 2–5 July 2019, Montréal, Quebec, Canada (pp. 99–108). International Educational Data Mining Society.

Karumbaiah, S., Ocumpaugh, J., Labrum, M., & Baker, R. S. (2019). Temporally rich features capture variable performance associated with elementary students' lower math self-concept. Paper presented at the Workshop on Online Learning and Social-Emotional Learning at the 9th International Conference on Learning Analytics and Knowledge (LAK '19), 4–8 March 2019, Tempe, Arizona, USA. New York: ACM.

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.

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

https://doi.org/10.1007/s40593-014-0019-7

Knight, S., Buckingham Shum, S., & Littleton, K. (2013). Epistemology, pedagogy, assessment and learning analytics. *Proceedings of the 3rd International Conference on Learning Analytics and Knowledge* (LAK '13), 8–12 April 2013, Leuven, Belgium (pp. 75–84). New York: ACM. https://doi.org/10.1145/2460296.2460312

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Kyle, K., & Crossley, S. A. (2017). Assessing syntactic sophistication in L2 writing: A usage-based approach. *Language Testing, 34*(4), 513–535. https://doi.org/10.1177/0265532217712554

Kyle, K., Crossley, S. A., & Berger, C. (2018). The tool for the automatic analysis of lexical sophistication version 2.0. *Behavior Research Methods, 50*(3), 1030–1046. https://doi.org/10.3758/s13428-017-0924-4

Lust, G., Elen, J., & Clarebout, G. (2013). Students' tool-use within a web-enhanced course: Explanatory mechanisms of students' tool-use pattern. *Computers in Human Behavior, 29*(5), 2013–2021. https://doi.org/10.1016/j.chb.2013.03.014

Marsh, H. W., Parker, J. W., & Smith, I. D. (1983). Preadolescent self-concept: Its relation to self-concept as inferred by teachers and to academic ability. *British Journal of Educational Psychology, 53*(1), 60–78. https://doi.org/10.1111/j.2044-8279.1983.tb02536.x

Marsh, H. W., Smith, I. D., & Barnes, J. (1985). Multidimensional self-concepts: Relations with sex and academic achievement. *Journal of Educational Psychology, 77*(5), 581–596. https://doi.org/10.1037/0022-0663.77.5.581

McQuiggan, S. W., Mott, B. W., & Lester, J. C. (2008). Modeling self-efficacy in intelligent tutoring systems: An inductive approach. *User Modeling and User-Adapted Interaction, 18*(1–2), 81–123. https://doi.org/10.1007/s11257-007-9040-y

Miller, G. A. (1995). WordNet: A lexical database for English. *Communications of the ACM, 38*(11), 39–41. https://doi.org/10.1145/219717.219748

Miller, W., Baker, R., Labrum, M., Petsche, K., Liu, Y.-H., & Wagner, A. (2015). Automated detection of proactive remediation by teachers in reasoning mind classrooms. *Proceedings of 5th International Learning Analytics and Knowledge Conference* (LAK '15), 16–20 March 2015, Poughkeepsie, NY, USA (pp. 290–294). New York: ACM. https://doi.org/10.1145/2723576.2723607

Mingle, L. (2013). *Threats to Success in Mathematics: Examining the Combined Effects of Choking Under Pressure and Stereotype Threat* (Doctoral dissertation, University of Illinois at Urbana-Champaign).

Moon, S., Potdar, S., & Martin, L. (2014). Identifying student leaders from MOOC discussion forums through language influence. *Proceedings of the Workshop on Modeling Large Scale Social Interaction in MOOCs at the 2014 Conference on Empirical Methods in Natural Language Processing* (EMNLP 2014), 25 October 2014, Doha, Qatar (pp. 15–20). Association for Computational Linguistics. https://doi.org/10.3115/v1/w14-4103

Nasir, N. I. S., & Saxe, G. B. (2003). Ethnic and academic identities: A cultural practice perspective on emerging tensions and their management in the lives of minority students. *Educational Researcher*, *32*(5), 14–18. https://doi.org/10.3102/0013189x032005014

Ocumpaugh, J., San Pedro, M. O., Lai, H. Y., Baker, R. S., & Borgen, F. (2016). Middle school engagement with mathematics software and later interest and self-efficacy for STEM careers. *Journal of Science Education and Technology, 25*(6), 877–887. https://doi.org/10.1007/s10956-016-9637-1

Osborne, J. W., & Jones, B. D. (2011). Identification with academics and motivation to achieve in school: How the structure of the self influences academic outcomes. *Educational Psychology Review, 23*(1), 131–158. https://doi.org/10.1007/s10648-011-9151-1

Osterman, K. F. (2000). Students' need for belonging in the school community. *Review of Educational Research, 70*(3), 323–367. https://doi.org/10.3102/00346543070003323

Pajares, F., & Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology, 86*(2), 193–203. https://doi.org/10.1037/0022-0663.86.2.193

Pardo, A., Han, F., & Ellis, R. A. (2016). Exploring the relation between self-regulation, online activities, and academic performance: A case study. *Proceedings of the 6th International Conference on Learning Analytics and Knowledge* (LAK '16), 25–29 April 2016, Edinburgh, UK (pp. 422–429). New York: ACM. https://doi.org/10.1145/2883851.2883883

Prenzel, M. (1992). The selective persistence of interest. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 71–98). Hillsdale, NJ: Lawrence Erlbaum. https://doi.org/10.1002/acp.2350080510

R Core Team. (2018). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Renninger, K. A. (2009). Interest and identity development in instruction: An inductive model. *Educational*

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

Psychologist, 44(2), 105–118. https://doi.org/10.1080/00461520902832392

Reveles, J. M., Cordova, R., & Kelly, G. J. (2004). Science literacy and academic identity formulation. *Journal of Research in Science Teaching, 41*(10), 1111–1144. https://doi.org/10.1002/tea.20041

Roberts, B., & DelVecchio, W. (2000). The rank-order consistency of personality from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin, 126*(1), 3–25. https://doi.org/10.1037/0033-2909.126.1.3

Ryan, K., & Ryan, A. (2005). Psychological processes underlying stereotype threat and standardized math test performance. *Educational Psychologist, 40*(1), 53–63. https://doi.org/10.1207/s15326985ep4001_4

San Pedro, M. O., Ocumpaugh, J., Baker, R. S., & Heffernan, N. T. (2014). Predicting STEM and non-STEM college major enrollment from middle school interaction with mathematics educational software. In J. Stamper et al. (Eds.), *Proceedings of the 7th International Conference on Educational Data Mining* (EDM2014), 4–7 July 2014, London, UK (pp. 276–279). International Educational Data Mining Society.

Sansone, C., Weir, C., Harpster, L., & Morgan, C. (1992). Once a boring task always a boring task? Interest as a self-regulatory mechanism. *Journal of Personality and Social Psychology, 63*(3), 379–390. https://doi.org/10.1037/0022-3514.63.3.379

Schlenker, B., & Weigold, M. (1989). Goals and the self-identification process: Constructing desired identities. In L. A. Pervin (Ed.), *Goal concepts in personality and social psychology* (pp. 243–289). Hillside, NJ: Lawrence Erlbaum Association.

Shavelson, R., & Bolus, R. (1982). Self-concept: The interplay of theory and methods. *Journal of Educational Psychology, 74*(1), 3–17.

Sirbu, M. D., Dascalu, M., Crossley, S., McNamara, D. S., & Trausan-Matu, S. (2019). Longitudinal analysis of participation in online courses powered by cohesion network analysis. In K. Lund, G. Niccolai, E. Lavoué, C. Hmelo-Silver, G. Gweon, & M. Baker (Eds.), *A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings, Proceedings of the 13th International Conference on Computer Supported Collaborative Learning* (CSCL 2019), Volume 2, 17–21 June 2019, Lyon, France (pp. 640–643). International Society of the Learning Sciences. https://doi.org/10.1007/978-3-319-98572-5_59

Slater, S., Ocumpaugh, J., Baker, R., Lib, J., & Labrum, M. (2018). Identifying changes in math identity through adaptive learning systems use. In J. C. Yang, M. Chang, L.-H. Wong, & M. T. Rodrigo (Eds.), *Proceedings of the 26th International Conference on Computers in Education* (ICCE 2018), 26–30 November 2018, Manila, Philippines (pp. 71–76). Asia-Pacific Society for Computers in Education.

Solomon, Y. (2007). Not belonging? What makes a functional learner identity in undergraduate mathematics? *Studies in Higher Education, 32*(1), 79–96. https://doi.org/10.1080/03075070601099473

Steinmayr, R., & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning & Individual Differences, 19*(1), 80–90. https://doi.org/10.1016/j.lindif.2008.05.004

Stipek, D. J. (1981). Children's perceptions of their own and their classmates' ability. *Journal of Educational Psychology*, *73*(3), 404–410. https://doi.org/10.1037/0022-0663.73.3.404

Syed, M., Azmitia, M., & Cooper, C. R. (2011). Identity and academic success among underrepresented ethnic minorities: An interdisciplinary review and integration. *Journal of Social Issues, 67*(3), 442–468. https://doi.org/10.1111/j.1540-4560.2011.01709.x

Syed, M., & Chemers, M. M. (2011). Ethnic minorities and women in STEM: Casting a wide net to address a persistent social problem. *Journal of Social Issues*, *67*(3), 435–441. https://doi.org/10.1111/j.1540-4560.2011.01708.x

Walton, G. M., & Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science, 331*(6023), 1447–1451. https://doi.org/10.1126/science.1198364

Watt, H. M., Hyde, J. S., Petersen, J., Morris, Z. A., Rozek, C. S., & Harackiewicz, J. M. (2017). Mathematics: A critical filter for STEM-related career choices? A longitudinal examination among Australian and US adolescents. *Sex Roles*, *77*(3–4), 254–271. https://doi.org/10.1007/s11199-016-0711-1

Wen, M., Yang, D., & Rosé, C. P. (2014a). Sentiment analysis in MOOC discussion forums: What does it tell us? In J. Stamper et al. (Eds.), *Proceedings of the 7th International Conference on Educational Data Mining* (EDM2014), 4–7 July 2014, London, UK (pp. 130–137). International Educational Data Mining Society.

Wen, M., Yang, D., & Rosé, C. P. (2014b). Linguistic reflections of student engagement in massive open online courses. *Proceedings of the 8th International AAAI Conference on Weblogs and Social Media* (ICWSM '14), 1–4 June 2014, Ann Arbor, Michigan, USA (pp. 525–534). Palo Alto, CA: AAAI Press.

Winne, P. H., & Jamieson-Noel, D. (2002). Exploring students' calibration of self reports about study tactics and achievement. *Contemporary Educational Psychology, 27*(4),551–572. https://doi.org/10.1016/s0361-476x(02)00006-1

Zhou, M., & Winne, P. H. (2012). Modeling academic achievement by self-reported versus traced goal orientation. *Learning*

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

and Instruction, 22(6), 413–419. https://doi.org/10.1016/j.learninstruc.2012.03.004

Appendix: Math Identity Survey Questions

Self-Concept

1 How are you doing in math this year?
Bad Not good Good Great

2 Is math harder for you than subjects like reading or science?
A lot harder A little harder A little easier Much easier

3 Are you better at math than most of the other kids in your class?
Not as good Almost as good A little better A lot better

4 I have always been good at math.
Not true Somewhat true Mostly true Very true

5 I have always been smart.
Not true Somewhat true Mostly true Very true

Interest

6 Math is one of the most important subjects to me.
Not at all important Not that important Important Very Important

7 I find working on math assignments
Interesting Somewhat interesting Somewhat boring Boring

8 How much do you like math?
Hate Dislike Like Love

9 Compared to other things you learn in school, how much do you like math?
Much less A little less A little more A lot more

Value

10 How important is it to you to get good grades in math class?
Not at all important Not that important Important Very Important

11 In math, it is most important to me to learn the math my teacher teaches.
Not at all important Not that important Important Very Important

12 Doing well in math is important to me.
Not at all important Not that important Important Very Important

13 It is important to me to be able to solve math problems.
Not at all important Not that important Important Very Important

14 If I practice, I can change how good I am at math.
Not true Somewhat true Mostly true Very true

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)