A Data-driven Path Model of Student Attributes, Affect, and Engagement in a Computer-based Science Inquiry Microworld

Arnon Hershkovitz, Ryan S.J.d. Baker, Janice Gobert, Adam Nakama
Worcester Polytechnic Institute, Worcester, MA, USA
{arnonh,rsbaker,jgobert,nakama}@wpi.edu

Abstract: Work in recent years has shown that a student's goals, attitudes, and beliefs towards learning impact their level of engagement during learning, and that engagement during learning plays a key role in learning outcomes. In this paper, we investigate the mechanisms through which a student's goals, attitudes, and beliefs impact the student's engaged and disengaged behaviors. In particular, we study whether affect is a mediating variable between learner attributes and engagement/disengagement during learning. To this end, we conduct exploratory path analysis on data from middle school learners who were conducting inquiry in science a microworld. We find weak but significant relationships between variables related to attitudes and beliefs and variables related to affective states and engagement. We present a path model that highlights boredom as an important mediator between a tendency towards avoiding novelty and off-task behavior during learning.

Introduction

Engagement is a critical factor in accomplishing learning tasks, in both traditional classroom and computer-based learning (Corno & Mandinach, 1983; Pintrich & Schrauben, 1992). Engagement, the extent to which students are involved in learning tasks, is positively linked with desirable learning outcomes such as deep learning and grades (Skinner, Wellborn & Connell, 1990; Wigfield et al., 2008). Individual differences in engagement are often due to differences in students' goals, beliefs, and attitudes (e.g., Zimmerman & Bandura, 1994). Affect has been hypothesized to be a key potential mediator between these relationships (e.g., Linnenbrink, 2007).

In this paper, we investigate how learners' goals, attitudes, and beliefs influence their degree of engagement during learning, and how this process is mediated by affective states, using data from middle school students completing a computer-based science inquiry activity in the domain of physical science. Although in recent years, research has revealed relationships between student attributes, affective states, and engagement behaviors while using computer-based learning environments, past studies have typically focused only on small subsets of these constructs. We, on the other hand, examine the relationships between a larger set of variables, and we do it using an exploratory path analysis, a method for describing the direct and indirect dependencies among a set of variables (Land, 1969).

Engaged and Disengaged Behaviors

Engagement and disengagement have been characterized as "elusive variables" (Corno & Mandinach, 1993), due to the difficulty of defining these constructs precisely. Our approach towards operationalizing engagement and disengagement is to identify and measure specific behaviors that indicate engagement and/or disengagement. One such "disengaged behavior" is off-task behavior (Karweit & Slavin, 1982), where the student ceases to engage in the learning task, and instead engages in an unrelated activity. Off-task behavior has been shown to lead to poorer learning outcomes in educational software (Baker et al., 2004; Rowe et al., 2009). Another form of disengaged behavior frequently seen in educational software is gaming the system, a behavior in which students engage in systematic guessing or rapid help requests in order to obtain answers rather than thinking through the learning material (Baker et al., 2004). Gaming the system also has been shown to be negatively correlated to learning (Baker et al., 2004).

Student Attributes and Disengagement

There is increasing evidence that semi-stable student attributes such as poor goal orientation (including the goal of avoiding failure – Elliot, McGregor, & Gable, 1999, and the goal of avoiding work – Harackiewicz et al 2002), low self-efficacy (Zimmerman & Bandura, 1994), and low grit ("perseverance and passion for long-term goals" – Duckworth et al., 2007) are associated with disengagement (cf. Elliot, McGregor, & Gable, 1999; Harackiewicz et al., 2000; 2002; Baker et al., 2008). At the same time, there is increasing evidence that student behaviors of various types mediate the relationships between student attributes and learning (cf. Skinner, Wellborn & Connell, 1990; Middleton & Midgley, 1997; Elliot, McGregor, & Gable, 1999; Blackwell, Trzesniewski & Dweck, 2007). For instance, goal orientation was found to be related to cognitive engagement:

students with learning goals tend to be more engaged, and tend to achieve better learning outcomes than students with performance goals (e.g., Wolters & Pintrich, 1996; Dupeyart & Mariné, 2005). Another example is grit, which predicts gaming the system on material the student does not know, in turn leading to poorer learning outcomes (Baker et al., 2008).

Affect and Disengagement

Another factor which plays an important role in student learning and engagement is affect (Craig et al., 2004), the subjective perception of emotional states (Corno, 1986) in context. The affective state of boredom has been shown to be negatively correlated with learning outcomes (Craig et al., 2004), while the affective state of "engaged concentration" (cf. Baker et al., 2010), also termed "flow", has been shown to be positively correlated with learning outcomes (Craig et al., 2004). The relationship between confusion and learning appears to be complex, with one study suggesting that the context of confusion determines its relationship to learning (Lee et al., 2011). In addition, affect has been found to be related to students' achievement goals and self-efficacy beliefs (e.g., Dweck & Leggett, 1988; Linnenbrink & Pintrich, 2002). Furthermore, it was shown that affect mediates the relationships between student goals and engagement, with pleasant affect having a positive mediating function and negative affect having a negative effect (Linnenbrink, 2007). To give another example, gaming the system is associated both with high levels of boredom (Baker et al., 2010) and low levels of grit (Baker et al., 2008), and boredom is positively associated with work avoidance (Duda & Nichols, 1992), a construct related to grit. This leads to the question: Does boredom potentially mediate the relationship between grit and gaming the system? And which affective states mediate the relationships between student attributes and other disengaged behaviors?

Methods

Participants

The study participants were 48 sixth-grade students (from 3 different classes), ranging in age from 10-12 years old, from a public middle school in a small city in Central Massachusetts. This school is majority White (68%), but has a substantial Hispanic minority (15%), and over 20% of residents in the city speak a language other than English at home. Overall, the city where the school is located has a median income only moderately lower than the U.S. median, but significantly lower than surrounding communities in the region, with almost 20% of individuals below the poverty level.

The relationship between carelessness and student goals was studied in the context of students' scientific inquiry within a physical science microworld. This microworld, for the domain of phase change, enabled students to explore how a particular substance, such as water, changes phases from solid to liquid to gas as it is heated. The goal of this environment and its associated activities was for students to engage in inquiry in order to develop an understanding that a substance has a melting point and a boiling point, which are both independent of the amount of sample. The phase change environment used, hosted by Science Assistments (www.scienceassistments.org; Gobert et al., 2009; Sao Pedro et al., in press) enabled students to engage in authentic inquiry using a microworld and inquiry support tools. Each task in the learning environment required students to conduct experiments to determine if a particular independent variable (container size, heat level, substance amount, or cover status) affected various outcomes (melting point, boiling point, time to melt, or time to boil).

Variables and Procedure

Within the study, we collected data relative to three categories of research variables, conceptualized as a 3-tier model. The three tiers studied are student attributes, affective states, and engaged and disengaged behaviors. We study the relationships between these tiers under the modeling assumption that the effect of student attributes on engagement is mediated by affective states. After receiving a short introduction to the activity, the students first completed the student attribute surveys online, and then engaged in the phase change learning activities for 20 minutes.

Student Attributes (11 variables): The variables are: *learning goal orientation*, the goal of developing skill or learning; *performance-approach goal orientation*, the goal of demonstrating competence; *performance-avoidance goal orientation*, the goal of avoiding demonstrating incompetence; *academic efficacy*; *avoiding novelty*; *disruptive behavior*; *self-presentation of low achievement*, the desire to prevent peers from knowing how well the student is performing; *skepticism about the relevance of school for future success*; *grit*, the perservance and passion for long-term goals; *work-avoidance*; and *self-efficacy for self-regulated learning*.

These variables are measured using surveys administered to students online. The Patterns of Adaptive Learning Scales (PALS) survey (Midgley et al., 1997) was used to assess semi-stable student attributes. In specific, two scales were utilized: PALS 1, Personal Achievement Goal Orientation (14 items), and PALS 4, Academic-related Perceptions, Beliefs, and Strategies (28 items), with all items in these scales rated on a 5-point Likert scale. Also, surveys were used to measure *grit* on a 5-point Likert scale (Duckworth et al., 2007), *work-avoidance* on a 7-point Likert scale (Harackiewicz et al., 2000), and *self-efficacy for self-regulated learning* on a 7-point Likert scale (Bandura, 1990).

Engaged and Disengaged Behaviors (4 variables): The following engaged and disengaged behaviors were coded: *on task*, *on-task conversation*, *off-task*, and *gaming the system*. These variables were assessed using quantitative field observations during student use of the learning environment (cf. Baker et al. 2004, 2010). For each student and for each observed variable, the percentage of the instances in which each behavior was observed was computed. Observed were conducted by trained observers (in specific, the second and fourth authors), who had previously achieved inter-rater reliability over 0.6 for the coding protocol and both of the coding schemes used in this study, with a different population. This level of inter-rater reliability, while imperfect, is substantially higher than that seen for expert coding of academic emotions and behaviors of this nature in video data (cf. Graesser et al., 2006). Each observation lasted twenty seconds, and was conducted using peripheral vision. That is, the observers stood diagonally behind or in front of the student being observed and avoided looking at the student directly, in order to make it less clear when an observation was occurring (cf. Baker et al. 2004, 2010). If two distinct behaviors were seen during an observation, only the first behavior observed was coded; similarly, if two distinct affective states were seen during an observation, only the first was coded. Only the student currently being observed was coded during each observation period, to avoid bias towards more salient behaviors and affective states. Observations were recorded using an app for the Google Android platform, which enforced the coding procedure and coding schemes. For each student and for each variable, the percentage of the instances in which this variable was observed was computed.

Affective States (4 variables): The following affective states were measured: *confusion*, *engaged concentration*, *boredom*, and *frustration*. These observations were conducted concurrently to the observations of engaged and disengaged behaviors, using the same procedure and coding schemes as in Baker et al. (2010). For each student and for each observed variable, the percentage of the instances in which each affective state was observed were computed. During analysis, the affective state of engaged concentration was treated as representing a different state when students were off-task as opposed to on-task, as being deeply engaged in off-task behavior should not be considered evidence for engagement with the learning material.

Analysis Approach: Path Models

Path models are an extension of multiple regression, in which the relationships between multiple variables are evaluated simultaneously. Path models are a graphical representation of a Structural Equation Model (SEM) describing linear relationships between the variables under consideration (Land, 1969). In recent years, path models have been used to shed light on the relationships between multiple learning variables. In most cases, path models have served as a means to test existing theoretical models of the relationships between variables of different types, e.g., goals, metacognitive, study strategies, and academic achievement (Dupeyart & Marine, 2005; Blackwell, Trzesniewski, & Dweck, 2007). However, in other cases, path models have been built bottom-up, to find the best possible model for a given dataset without assuming the model a-priori (cf. Fouad & Smith, 1996; Iverach & Fisher, 2008). In this paper, given the lack of a prior theoretical model linking all of the constructs under investigation, we adopt a bottom-up method, essentially using path models as an educational data mining method (cf. Baker & Yacef, 2009).

To this end, we follow a heuristic inspired by Cohen et al. (1993) for building path models in this fashion: potential links are chosen to be added to the path model according to the degree of improvement in the model's fit to the data. We impose only one type of structure: a 3-tier model, consisting of the following sets of variables: student attributes (tier one), affective states (tier two), and behaviors indicating engagement or disengagement (tier three). Attributes are assumed to influence later affective states, and both attributes and affective states are assumed to influence behavior. The model creation process was initiated using a set of models, each of which consisted of only two variables and one link, based on the set of correlations which were statistically significant under post-hoc correction (listed below). Then, we assessed the impact on model significance from adding additional variables and links to the model. Models were built and assessed using the AMOS 19 software for path modeling (Arbuckle, 2010).

Results

First, in order to understand the frequencies of the research variables, we computed means and standard deviations for each of the research variables. A full list of proportions, means and standard deviations of the variables used is given in Table 1. The average proportion of off-task behavior (13%) is within the range previously observed in studies of students using intelligent tutor software in the U.S. (cf. Baker et al., 2004;

Baker, 2009; Cetintas et al., 2010), and is similar or slightly lower than proportions of off-task behavior seen in traditional classrooms (cf. Karweit & Slavin, 1982).

The proportion of gaming the system (1%) is substantially lower than seen in previous studies of intelligent tutoring systems and educational games (Baker et al., 2004, 2010; Walonoski & Heffernan, 2006). Students were observed to be in a state of engaged concentration 59% of the time, a proportion consistent with past research on students using intelligent tutoring systems (Craig et al., 2004; Baker et al., 2010), though higher than seen in some studies (Sabourin et al., 2011). Boredom was substantially more common than in several of these past studies, occurring 20% of the time, while confusion was relatively less common, occurring 8% of the time.

Table 1: Mean and standard deviation for the research variables.

Student Attributes (all but the last two are on a 5-Likert scale; the last two are on a 7-Likert scale)			**Affective States (proportion of observed instances to total number of observations), N=48**		
Variable	**Mean**	**SD**	**Variable**	**Mean**	**SD**
learning goal orientation (N=46)	4.20	0.59	*frustration*	0.06	0.06
performance-approach goal orientation (N=46)	3.00	1.19	*confusion*	0.08	0.07
performance-avoid goal orientation (N=46)	3.09	1.24	*engaged concentration*	0.59	0.18
academic efficacy (N=46)	3.57	0.90	*boredom*	0.20	0.13
avoiding novelty (N=46)	2.90	1.10	**Engagement (proportion of observed instances to total number of observations), N=48**		
disruptive behavior (N=46)	2.77	1.21			
self-presentation of low achievement (N=45)	2.14	1.00	**Variable**	**Mean**	**SD**
skepticism about the relevance of school for future success (N=45)	2.65	1.12	*on task*	0.75	0.16
grit (N=43)	2.86	0.65	*on-task conversation*	0.11	0.10
work-avoidance (7-Likert scale) (N=47)	4.06	1.19	*off-task*	0.13	0.11
self-efficacy for self-regulated learning (7-Likert scale) (N=47)	4.52	0.88	*gaming the system*	0.01	0.04

Relationships between the Research Variables

In order to explore the relationships between the research variables, we examined the inter-correlations between the variables of the different tiers. Results are summarized in Table 2. As we ran an overall of 104 bivariate correlations tests, we corrected for multiple comparisons using the post-hoc False Discovery Rate (FDR) method, which produces a q-value that can be interpreted the same way as a p-value. Between student attributes and affective states, only one pair of variables (out of 44) was found to be significantly correlated when controlling for multiple comparisons: *disruptive behavior* and *confusion* ($r=0.45$, $q<0.05$). This relationship might be explained as the following: as disruptive behavior increases, less learning opportunities are available for the student, hence she or he is more likely to be confused when required to demonstrate their skills. It is surprising that none of the three main goal orientation scales (learning, performance-approach, and performance-avoid) were found to be significantly correlated with any of the affective states, considering previous studies (e.g., Linnenbrink & Pintrich, 2002).

Between student attributes and engagement tiers, no significant correlations were found. Again, no correlations were found between the three main goal orientation scales with any of the engagement variables. The strongest correlation (in terms of q-value) was between *performance-approach goal orientation* and *on task conversation* ($r=0.20$, $q=0.73$). This is consistent with previous studies which have found no relationships between goal orientation and disengaged behaviors within students using educational software (e.g. Baker, 2007; Baker et al., 2008, Rowe et al., 2009).

Finally, between the affective states and engaged/disengaged behaviors, four significant correlations were found. *Engaged concentration* was found to be significantly correlated with *on-task* ($r=0.65$, $q<0.01$) and *off-task* (-0.54, $q<0.01$); also, *boredom* was found to be significantly correlated with *on-task* ($r=0.57$, $q<0.01$) and *off-task* ($r=0.65$, $q<0.01$).

One complication in analyzing these relationships is that the affective states and engagement variables are not normally distributed. Testing for multivariate normality, we used Small's test for skewness (Small, 1980) and Anscome & Glynn's (1983) variant of Small's test for kurtosis, and found significant evidence for skew, Q_1=71.28 (df=19), p<0.0001, and for multivariate kurtosis, Q_2=56.74 (df=19), at p<0.0001. Hence, nonparametric bootstrapping and Maximum Likelihood estimation were used to generate model parameters and estimates (cf. Kline, 2010), with removal of cases with missing values during bootstrapping.

Table 2: Inter-correlations between the research variables of different tiers (significant values after post-hoc corrections bolded).

Variables		**Affective States**				**Engagement**			
		frust.	*conf.*	*concent.*	*boredom*	*on task*	*on-task conv.*	*off-task*	*gaming*
Student Attributes	*learning goal* (N=46)	-0.16	0.04	0.06	0.09	-0.02	0.00	0.03	0.00
	perf-approach.goal (N=46)	-0.07	0.26	0.12	-0.25	-0.01	0.20	-0.12	-0.13
	perf.-avoid goal (N=46)	-0.16	-0.01	0.11	-0.15	0.07	0.09	-0.15	-0.09
	academic. efficacy (N=46)	-0.38	-0.15	0.32	-0.17	0.15	0.00	-0.11	-0.24
	avoiding. novelty (N=46)	0.01	0.13	0.21	-0.35	0.31	0.13	-0.34	0.10
	disruptive behavior (N=46)	0.00	**0.45***	-0.12	-0.10	-0.16	0.23	0.03	-0.05
	self-presnt. of low achv. (N=46)	-0.13	0.22	0.07	-0.19	0.11	0.00	-0.15	-0.03
	skepticism (N=45)	-0.09	0.34	0.07	-0.14	0.12	0.06	-0.23	0.01
	grit (N=43)	0.17	-0.03	0.07	-0.04	0.24	0.00	-0.28	-0.14
	work-avoidance (N=47)	0.05	0.11	-0.05	-0.02	-0.08	0.24	-0.14	-0.01
	self-efficacy (N=47)	-0.16	-0.12	0.19	-0.14	-0.06	0.04	-0.01	0.00
Affective States	*frustration* (N=46)					-0.19	0.24	-0.04	0.23
	confusion (N=46)					-0.02	0.14	-0.09	-0.01
	concentration (N=46)					**0.65****	-0.34	**-0.54****	-0.09
	boredom (N=46)					**-0.57****	0.18	**0.65****	-0.11

* q<0.05, ** q<0.01

Path Model

As was previously mentioned, we applied a bottom-up construction heuristic for the path model. Figure 1 presents the best fitting model found, involving only three variables, one of each tier: *avoiding novelty*, *boredom*, and *off-task*. The standardized regression coefficient between *avoiding novelty* and *boredom* is -0.28, which is marginally statistically significant, with p=0.06; the coefficient between *boredom* and *off-task* is 0.48, which is statistically significant at p<0.01. This model was validated using chi-square, obtaining a value of 1.24 (df=1), p=0.27 (This test compares the best model to a full (saturated) model; a non-significant chi-square is therefore desirable). For this test, a ratio of chi-square to df which is equal to or smaller than 3 is considered acceptable (Schreiber et al., 2006). Three measure of goodness of fit, used for small samples that are not multivariate normal, CFI (comparative fit index), GFI (goodness of fit index), and IFI (incremental fit index), got values of 0.98 each; the agreed cutoff for these measures is 0.95.

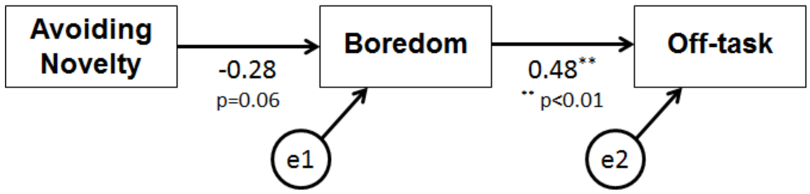

Figure 1. The final path model: Rectangles represent observed variables, circles – latent variables

This model suggests that the negative relationship between *avoiding novelty* and *off-task* behavior is mediated by *boredom*. In other words, students who avoid novelty are less likely to be bored during learning, and students who are less bored are less likely to become off-task. We can also understand this model as saying that students who desire novelty are more likely to be bored and therefore are more likely to be off-task. The relationship between avoiding novelty and off-task behavior is much more understandable when the mediating factor of boredom is taken into account.

Discussion & Conclusions

Within this paper we have analyzed the relationships between three categories of variables related to students' characteristics and behavior: a) student attributes, including goal orientation, beliefs regarding learning, and attitudes towards learning; b) affective states during learning; and c) engagement and disengagement during learning. Overall, there were only a few statistically significant correlations in between the different categories. Interestingly, this study failed to replicate previous results that establish relationships between goal orientation and engagement (e.g., Dweck & Leggett, 1988; Elliot & Harackiewicz, 1996), as well as between goal orientation and affective states (e.g., Pekrun, Elliot, & Maier, 2006). A possible explanation is that the domain plays a major role in the lack of these relationships, as inquiry-based learning environments foster better engagement and may lead to a greater probability of mastery goal orientation during learning (cf. Hmelo-Silver, Duncan & Chinn, 2007), hence changing the nature of interactions between variables. Indeed, previously there were no associations found between carelessness (a different measure of disengagement) and goal orientation (Hershkovitz et al., 2011), and no relationships between goal orientation and achievement (Shimoda, White & Frederiksen, 2002) in science inquiry tasks. In general, goal orientation has been a poor predictor of engagement in educational software (Baker, 2007; Baker et al., 2008; Rowe et al., 2009), despite its general importance in other contexts (Wolters & Pintrich, 1996; Dupeyart & Mariné, 2005; Blackwell et al., 2007).

We applied a data-driven approach for constructing a path model to describe the relationships between the research variables. Avoiding novelty was found to be related to boredom, which is in turn related to off-task behavior. The model suggests that it will be valuable to proactively scaffold students who tend to avoid novelty in ways that promote interest. It is important to emphasize that the path model by itself does not imply causality, as is generally the case for post-hoc correlational analyses. In general, it will be valuable to replicate this analysis using different datasets from the same domain, and datasets from additional domains, in order to determine how general the patterns are.

One other surprising finding was the relative rarity of gaming the system compared to prior research in intelligent tutoring systems and educational games. One explanation is that gaming the system is more difficult in the current learning environment, given the relatively lower degree of feedback than seen in these types of environments. Interestingly, carelessness – another behavior indicating disengagement – has also been found to be less common in Science Assistments than in mathematics tutoring systems (Hershkovitz et al., 2011).

In summary, the research presented here suggests that in the context of computer-based science inquiry, there is a substantial relationship between avoiding novelty (a student attribute) and off-task behavior (a disengaged behavior), mediated by boredom (an affective state). In general, it will be valuable to study further how and why boredom mediates this relationship. One method for doing so will be to apply temporal analysis, exploring how engagement and affect change during the learning task.

References

Anscombe, F.J. & Glynn, W.J. (1983). Distribution of kurtosis statistic b2 for normal statistics. *Biometrika, 70*, 227–234.

Arbuckle, J.L. (2010). *IBM SPSS Amos 19 User's Guide*. Chicago, IL: IBM.

Baker, R.S.J.d. (2007). Modeling and understanding students' off-task behavior in Intelligent Tutoring Systems. *Proceedings of ACM CHI 2007: Computer-Human Interaction*, 1059-1068.

Baker, R.S.J.d. (2009) Differences between intelligent tutor lessons, and the choice to go off-task. *Proceedings of the 2nd International Conference on Educational Data Mining*, 11-20.

Baker, R.S.J.d., Corbett, A.T., Koedinger, K.R., Wagner, A.Z. (2004) Off-task behavior in the cognitive tutor classroom: When students "game the system". *Proceedings of ACM CHI 2004: Computer-Human Interaction*, 383-390.

Baker, R.S.J.d., D'Mello, S.K., Rodrigo, M.M.T., Graesser, A.C. (2010) Better to Be Frustrated than Bored: The Incidence, Persistence, and Impact of Learners' Cognitive-Affective States during Interactions with Three Different Computer-Based Learning Environments. *International Journal of Human-Computer Studies, 68*(4), 223-241.

Baker, R.S.J.d., Walonoski, J., Heffernan, N., Roll, I., Corbett, A., & Koedinger, K. (2008). Why students engage in "gaming the system" behavior in interactive learning environments. *Journal of Interactive Learning Research, 19*(2), 185-224.

Baker, R.S.J.d., Yacef, K. (2009) The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining, 1*(1), 3-17.

Bandura, A. (1990). Multidimensional Scales of Perceived Self-efficacy. Stanford, CA: Stanford University.

Blackwell, L.S., Trzesniewski, K.H., & Dweck, C.S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78*(1), 246-263.

Cetintas, S., Si, L., Xin, Y.P., & Hord, C. (2010). Automatic detection of off-task behaviors in intelligent tutoring systems with machine learning techniques. *IEEE Transactions on Learning Technologies, 3*(3), 228-236.

Cohen, P.R., Ballesteros, L., Carlson, A., & Amant, R.St. (1993). Automating path analysis for building causal models from data: First results and open problems. In *Proceedings of the Tenth International Conference on Machine Learning*, 57-64.

Corno, L. (1986). The metacognition control components of self-regulated learning. *Contemporary Educational Psychology, 11*, 333-346.

Corno, L. & Mandinach, E.B. (1983). The role of cognitive engagement in classroom learning and motivation. *Educational Psychologist, 18*(2), 88-108.

Craig, S.D., Graesser, A.C., Sullins, J., & Gholson, B. (2004). Affect and learning: An exploratory look into the role of affect in learning with AutoTutor. *Journal of Educational Media, 29*(3), 241-250.

Duckworth, A.L., Peterson, C., Matthews, M.D., & Kelly, D.R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and social Psychology, 92*(6), 1087-1101.

Duda, J., & Nicholls, J. (1992). Dimensions of achievement motivation in school work and sport. *Journal of Educational Psychology*, 84, 290-299.

Dupeyrat, C. and 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.

Dweck, C.S. and Leggett, E.L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*, 256–273.

Elliot, A.J., & Harackiewicz, J.M. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology, 70*(3), 461−475

Elliot, A.J., McGregor, H.A., & Gable, S.L. (1999). Achievement goals, study strategies, and exam performance: A mediational analysis. *Journal of Educational Psychology*, *76*, 628–644.

Fouad, N.A. & Smith, P.L. (1996). A test of a social cognitive model for middle school students: Math and science. *Journal of Counseling Psychology, 43*(3), 338-346.

Gobert, J., Heffernan, N., Koedinger, K.., & Beck, J. (2009). *ASSISTments Meets Science Learning (AMSL).* Proposal (R305A090170) funded by the U.S. Department of Education.

Graesser, A.C., McDaniel, B., Chipman, P., Witherspoon, A., D'Mello, S., & Gholson, B. (2006). Detection of emotions during learning with AutoTutor. *Proceedings of the 28th Annual Meeting of the Cognitive Science Society,* 285-290.

Harackiewicz, J.M., Barron, K.E., Tauer, J.M., Carter, S.M., & Elliott, A.J. (2000). Short-term and long-term consequences of achievement goals: Predicting interest and performance over time. *Journal of Educational Psychology, 92*(2), 316-330.

Harackiewicz, J.M., Barron, K.E., Tauer, J.M., & Elliot, A.J. (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of Interest and performance from freshman year through graduation. *Journal of Educational Psychology, 94,* 562–575.

Hershkovitz, A., Wixon, M., Baker, R.S.J.d., Gobert, J., Sao Pedro, M. (2011) Carelessness and goal orientation in a science microworld. Poster paper. *Proceedings of the 15th International Conference on Artificial Intelligence in Education*, 462-465.

Hmelo-Silver, C.E., Duncan, R.G., & Chinn, C.A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist, 42*(2), 99-107.

Iverach, M.R. & Fisher, D.L. (2008). An interdisciplinary investigation of high schools' approaches to learning sciences: The relations amongst achievement goals, constructivist pedagogical dimensions, motivational beliefs and self-regulated learning. *Proceedings of the Fifth International Conference on Science, Mathematics and Technology Education*, 233-255.

Karweit, N. & Slavin, R.E. (1982). Time-on-task: Issues of timing, sampling and definition. Journal of Educational Psychology, 74, 844-851.

Kline, R.B. (2010). *Principles and Practice of Structural Equation Modeling (3rd edition)*. New York, NY: The Guilford Press.

Land, K.C., (1969). Principles of path analysis. *Sociological Methodology, 1*, 3-37.

Lee, D.M.C., Rodrigo, Ma.M.T., Baker, R.S.J.d., Sugay, J.O., & Coronel, A. (2011). Exploring the relationship between novice programmer confusion and achievement. In *Proceedings of the 4th Bi-annual International Conference on Affective Computing and Intelligent Interaction*.

Linnenbrink, E.A. (2007). The role of affect in student learning: A multi-dimensional approach to considering the interaction of affect, motivation, and engagement. In P.A. & P. Reinhard (eds.), *Emotion in Education (pp. 107-124).* San Diego, CA: Elsevier Academic Press.

Linnenbrink, E.A. & Pintrich, P.R. (2002). Achievement goals theory and affect: An asymmetrical bidirectional model. Educational Psychologist, 27(2), 69-78.

Middleton, M., & Midgley, C. (1997). Avoiding the demonstration of lack of ability: An underexplored aspect of goal theory. *Journal of Educational Psychology, 89*, 710–718

Midgley, C., Maehr, M., Hicks, L., Roeser, R., Urdan, T., Anderman, E., & Kaplan, A., Arun-Kumar, R., Middleton, M. (1997). *Patterns of Adaptive Learning Survey (PALS)*. Ann Arbor, MI: University of Michigan.

Pekrun, R., Elliot, A.J., & Maier, M.A. (2006). Acheivement goals and discrete achievement emotions: A theoretical model and prospective tests. *Journal of Educational Psychology, 98*(3), 583-597.

Pintrich, P.R. & Schrauben, B. (1992). Students' motivational beliefs and their cognitive engagement in classroom academic tasks. In D.H. Schunk & Meece, J.L. (eds.), *Student Perceptions in the Classroom (pp. 149-183)*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Rowe, J.P., McQuiggan, S.W., Robison, J.L., & Lester, J.C. (2009). Off-task behavior in narrative-centered learning environments. *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Education*, 99-106.

Sabourin, J., Mott, B., Lester, J. (2011) Modeling Learner Affect with Theoretically Grounded Dynamic Bayesian Network. *Proceedings of 4th International Conference on Affective Computing and Intelligent Interaction.*

Sao Pedro, M., Baker, R., Gobert, J., Montalvo, O., & Nakama, A. (in press). Using Machine-Learned Detectors of Systematic Inquiry Behavior to Predict Gains in Inquiry Skills. To appear in User Modeling and User-Adapted Interaction.

Schreiber, J.B., Stage, F.K., King, J., Nora, A., & Barlow, E.A. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research, 99*(6), 323-337.

Shimoda, T.A., White, B.Y., & Frederiksen, J.R. (2002). Student goal orientation in learning inquiry skills with modifiable software advisors. *Science Education, 86*(2), 244-263.

Skinner, E.A., Wellborn, J.G., & Connell, J.P. (1990). What it takes to do well in school and whether I've got it: A process model of perceived control and children's engagement and achievement in school. Journal of Educational Psychology, 82(1), 22-32.

Small, N.J.H. (1980). Marginal skewness and kurtosis in testing multivariate normality. *Applied Statistics, 29*, 85–87.

Walonoski, J. & Heffernan, N.T. (2006). Detection and analysis of off-task gaming behavior in intelligent tutoring systems. *Proceedings of the Eight International Conference on Intelligent Tutoring Systems*, 382-391.

Wigfield, A., Guthrie, J.T., Perencevich, K.C., Taboada, A., Klauda, S.L., McRae, A., & Barbosa, P. (2008). Role of reading engagement in mediating effects of reading comprehension instruction on reading outcomes. Psychology in the Schools, 45(5), 432-445.

Wolters, C.W., Yu, S.L., Pintrich, P.R. (1996). The relation between goal orientation and students' motivational beliefs and self-regulated learning. *Learning and Individual Differences, 8*(3), 211-238.

Zimmerman, B.J. & Bandura, A. (1994). Impact of self-regulatory influences on writing course attainment. *American Educational Research Journal, 31*(4), 845-862.

Acknowledgments

This research was supported by NSF grant DRL-1008649, awarded to Janice Gobert and Ryan Baker. We would also like to thank Michael Sao Pedro, Ermal Toto, Orlando Montalvo, and Juelaila Raziuddin for their role in the development of the physical science inquiry microworld.