

# The Influence of Learner Characteristics on Conducting Scientific Inquiry Within Microworlds

Janice D. Gobert  
Juelaila J. Raziuddin  
Michael A. Sao Pedro

Learning Sciences and Technologies Program, Worcester Polytechnic Institute, 100 Institute Road  
Worcester, MA 01609 USA

## Abstract

In this study, we address whether learner characteristics can provide data to inform adaptive scaffolding of scientific inquiry skills in our learning environment, Science Assistsments. We found that *academic efficacy* positively predicted students' skills at generating hypotheses; another subscale, *skeptical of school relevance*, negatively predicted students' skills at conducting controlled experiments, specifically controlling for variables strategies (CVS).

**Keywords:** science learning environment; log files; inquiry learning; learner characteristics

## Introduction

Social cognitive theory has long recognized the inter-relationships between the learner, the learning environment, and the learner's behaviors, and in recent years, understanding these relationships has become an important area of research (Crippen et al., 2009; Payne et al., 2007). With the proliferation of science learning environments, it is important to unpack the relationship between learner characteristics and learning processes in these environments. This is important because current conceptualizations of knowledge ontologies for science stress content, skills, and nature of science (Perkins, 1986), but these alone are not sufficient to account for the wide range of scores exhibited in achievement outcomes (Gobert & Baker, 2010). Additionally, developing empirical models of these relationships remains a challenge and thus, there is, to date, no integrated theoretical model. The results of this are twofold: 1) instructional designers are left with minimal data with which to address learner characteristics, and 2), providing adaptive support to students for conducting inquiry rich science learning environments on the basis of learner characteristics is nearly impossible. In this study, we address whether learner characteristics might provide data to inform scaffolding of scientific inquiry skills in our learning environment, Science Assistsments (Gobert et al., 2007, 2009, 2010). In terms of inquiry skills, we measure hypothesizing, controlling for variables (CVS), interpreting data, and communicating findings. We use a variety of measures including log files, which capture student's interactions within the learning environment, as well as open-format explanations. Our study provides a

methodological advantage over many in that we do not rely on self-report measures of cognitive processing as this presents barriers to face validity (Crippen et al., 2009). Rather, we use students' log files as indices of students' strategy use (Winne et al., 2000); log files have been shown to be valid measure of processing quality (Sins et al., 2007).

We hypothesize that during science inquiry within an educational learning environment, students' goal orientation and self-efficacy may influence the selection of strategies for inquiry. Some key findings in accordance with our hypothesis are as follows. Mastery learning goals (goal of deep learning) are positively correlated with deep processing (cf. Elliot et al., 1999; Sins et al., 2007) and are a good predictor of achievement (Wolters, 2004). Performance-approach goals (goal to demonstrate good performance) are associated with high grades and exam performance (Harackiewicz et al., 2000; Pintrich, 2000; Wolters, 2004). Performance-avoidance goals (goal to avoid poor performance) have been shown to be associated with poorer learning (Skaalvik, 1997) and self-handicapping behaviors (Elliot, et al., 1999; Urdan, 2004). High self-efficacy (belief that one can achieve the task at hand) is associated with deeper processing of material (Bandura, 1997; Pintrich, 1999), including in science microworlds (Sins et al., 2007) and has been found to positively predict student pre-post content learning and science inquiry skills (Nelson & Ketelhut, 2008).

## Methodology

### Participants

Participants were 70 eighth grade students, ranging in age from 12-14 years, from a public middle school in Central Massachusetts. Students belonged to one of six class sections and had one of two science teachers.

### Materials

Science Assistsments Learning Environment ([www.scienceassistsments.org](http://www.scienceassistsments.org)) is a learning environment for Physics, Life Science, and Earth Science that supports students to conduct scientific inquiry with microworlds.

*Density Microworlds.* Our "Mass and Density" microworlds, focus on the Massachusetts curricular framework's "properties of matter" learning strand for mass, volume, and density. We have two microworlds in which

the students explore the relationships between mass, volume, and density. Microworld one allows the student to change the type of liquid and container shape while measuring the weight and volume (see Figure 1). Microworld two (Figure 2) is based on Archimedes' principle of buoyancy, and introduces students to the notion of density as it relates to mass and volume.

*Content and Inquiry Assessments.* We developed two short standardized-test style assessments to baseline student

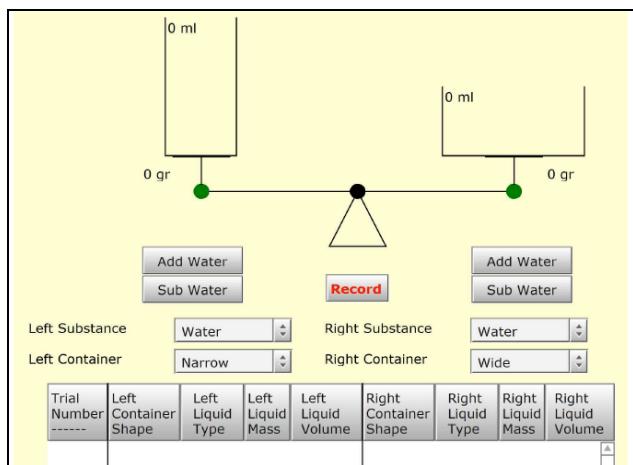


Figure 1: Microworld One - Balancing

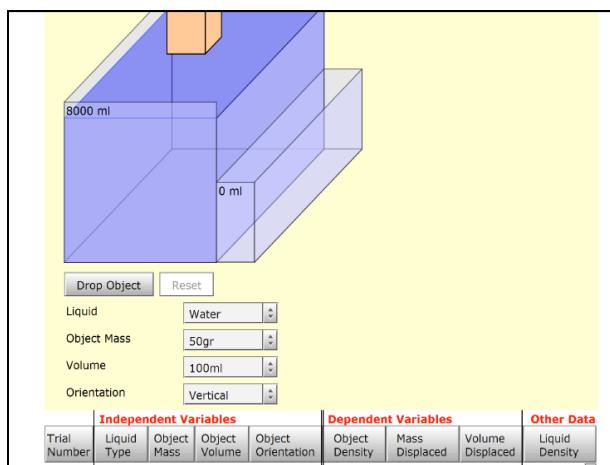


Figure 2: Microworld two - Archimedes

inquiry and density content knowledge. Our 13-item multiple choice inquiry test assesses students' understanding of hypotheses, designing controlled experiments, and analyzing data. Some items were developed by our team and others were acquired from Strand-Cary and Klahr (2009). Our density content test items assess students on density concepts that can be learned through exploration with our microworlds.

*Learner Characteristic Surveys.* We administered several subscales of the Patterns of Adaptive Learning Survey (PALS; Midgley, et al., 2000) including mastery learning orientation, performance-approach orientation, performance-avoidance orientation, academic efficacy, novelty

avoidance, disruptive behavior, self-presentation of low achievement, and skepticism of school's relevance for future success. We also administered the Ketelhut Self Efficacy scale for science inquiry (Ketelhut, 2007).

## Procedure

On Day 1, each section of students was introduced to the Science Assistance System, created accounts, and took the PALS survey (Midgley, et al., 2000) and Ketelhut scientific inquiry self-efficacy survey (Ketelhut, 2007). The Assistance system randomly assigned students into two groups, either density first followed by phase change or vice versa within each class section. This paper includes students who were given our density microworlds; however, for this paper we include only data from the density activity (N=70). Students took our density content and inquiry tests followed by four density activities. These activities progressed as follows. In each of four activities, students were oriented to each type of inquiry task by the task's description; specifically they were asked to write hypotheses, design and conduct experiments to test their hypotheses, interpret data, and communicate findings. (More information is given in the Data Coding section). Finally, students answered identical density and inquiry items as an immediate posttest.

## Data Coding

Inquiry Pre and Post Tests & Domain Pre and Post Tests were autoscored by the Assistments system (Razzaq et al., 2005).

*Fine-Grained Code Scheme for Student Open Responses.* Students' open-response data for hypothesizing, data interpretation, and communicating were hand scored by two independent coders. The coding scheme measures understanding of variables and their relationships. The coding scheme and scoring is as follows:

1. Did the student give the correct independent variable (IV) for the task? (0=not correct, 1=identify the correct IV)
2. Did the student give the correct dependent variable (DV) for the task? (0=not correct, 1=identify the correct DV)
3. Did the student explain the effect of the independent variable on the dependent variable? (0=no explanation, 1=explain IV and DV relationship, 2=accurate and detailed explanation of IV-DV relationships, including relevant information about density).

For *hypothesizing* there were a total of four tasks, two for each microworld; a cumulative score of these was used for our analyses to measure students' skill at hypothesizing.

For *interpreting data* students were asked to draw conclusions from their trials. This provided an assessment of students' skills at interpreting the relationships between independent and dependent variables. There were a total of four data interpretation tasks; a cumulative score of these was used to measure students' skill at interpreting data, and

For *communicating science processes*, these data were derived both from their hypotheses and data interpretation tasks with regard to the depth of explanations between the independent and dependent variable. In total, there were

eight communication tasks; a total of four tasks from each series of hypotheses and data interpretation. A cumulative score was used to measure students' skill at communicating. With all open response activities taken together across the two microworlds, possible scores ranged from 0-16 per inquiry skill. Inter-reliability between two coders on the three scales resulted in  $r = 0.88$ ,  $r = 0.78$ , and  $r = 0.67$  for *hypothesizing, data interpretation, and communicating findings*, respectively.

For *controlling for variables (CVS)*, students' log files were hand-scored by manually coding student activity sequences, called clips, using text replay tagging of log files (Sao Pedro, Baker, Gobert, Montalvo, & Nakama & Gobert, 2010; Montalvo, Baker, Sao Pedro, Nakama & Gobert, 2010), an extension to the text replay approach developed in Baker, Corbett, and Wagner (2006). A text replay is a pre-specified chunk of student actions presented in text that includes information such as each student action's time, type, widget selection, and exact input. Text replay tagging enables us to label multiple behaviors or skills, such as controlling for variables and testing hypotheses, independently within log files at the same time. We chose this approach to code these skills because we believe we can rigorously capture the nature of these skills in an open-ended learning environment. For example, in the case of controlling for variables, others have measured this skill using the percentage of pairwise controlled experiments (McElhaney & Linn, 2010). In our learning environment, a student may run several repeated trials to observe the microworld, then change one variable and run several more repeated trials to observe again. Using text replay tagging, we would label such an action sequence as demonstrating CVS whereas the successive pairwise controlled experiments rule would yield a low CVS estimate.

Student activity sequences, or clips, were composed of fine-grained actions as students typed their hypotheses and collected data within the Archimedes density microworld tasks. Since there were three tasks, each student had three clips that were hand-scored. As part of this process, we defined 4 tags that corresponded to systematic and haphazard data collection behaviors of interest, any or all of which could be used to classify a clip. These were "No Trials", "Controlling for Variables Strategy (CVS) trials", "Identifying Independent Variables (IV)", "Without Thinking Fastidiously". Specific to our analyses, we tagged a clip as "CVS" if the clip contained actions indicative of designing and running controlled experiments. We tagged "Identifying Independent Variable" if the clip had actions indicating attempts to test the independent variable chosen to be tested, regardless of whether or not proper CVS procedure was used. Thus, we may label a students' experiment log as CVS-compliant, hypothesis testing-compliant, both, or neither.

To ensure we adequately captured the constructs, two human coders tagged a subset of the data collection clips to generate a corpus of hand-coded clips using Text Replay Tagging software (Sao Pedro, et al., 2010; Montalvo, et al.,

2010). The corpus contained 213 clips. The human coders both tagged the first 50 clips to test for agreement; one human coder coded all remaining clips. In line with our approach, a human coder chooses at least one but possibly several tags to classify the clip. Agreement for the 50 clips tagged by both coders was high overall. There was an average agreement of  $\kappa = 0.87$  over all ten tags. More specifically and of importance to this work, there was good agreement on the CVS and testing hypotheses tags,  $\kappa = .66$  and  $\kappa = .81$ , respectively. High Kappa values suggest good agreement between coders; this degree of agreement was achieved in part through extensive discussion and joint labeling prior to the inter-rater reliability session.

## Analyses and Results

### *Using Learner Characteristics to Predict Hypothesizing, Interpreting Data, and Communicating Scientific Processes.*

We analyzed if any student learner characteristic subscales (for learning orientation or self-efficacy) or pretests (content or inquiry) could predict students' degree of skill in the three scientific inquiry skills of interest.

As first step, we tested for correlations amongst all of our variables. We found: Academic efficacy correlated positively with hypothesizing skill ( $r = .32$ ,  $p < .004$ ), with data interpretation skill ( $r = .25$ ,  $p < .019$ ), and with communicating skill ( $r = .26$ ,  $p < .002$ ). Mastery learning orientation was positively correlated with both data interpretation skill ( $r = .25$ ,  $p < .017$ ) and hypothesizing skill ( $r = .29$ ,  $p < .008$ ). Less desirable self-reported learner characteristics like novelty avoidance, disruptive behavior, presenting oneself as a low achiever, and being skeptical of school's relevance to success were negatively correlated with the four inquiry skills: Self-reported disruptive behavior correlated most strongly with communicating science processes ( $r = -.29$ ,  $p < .008$ ) and with interpreting data ( $r = -.32$ ,  $p < .004$ ), both moderate correlations. Also, the content density pretest correlated more strongly with data interpretation and communicating science processes inquiry skills than general inquiry pretest ( $r = .610$ , and  $r = .634$ ) respectively. Both content density pretest and general inquiry pretest correlated strongly to hypothesizing skill ( $r = .549$ , and  $r = .548$ ).

As a second step in order to determine which factors best predicted skill in hypothesizing, interpreting data, and communicating, we performed three forward-selection regressions with each inquiry skill as a dependent measure. As shown in Table 1, most notably, the density content pretest appeared as a significant predictor of the three inquiry skills, and was the first variable entered. By itself, the content pre-test predicted  $R^2=30\%$  for hypothesizing ( $F(1,69) = 29.40$ ,  $p < .001$ ),  $R^2=37\%$  for interpreting data ( $F(1,69) = 40.37$ ,  $p < .001$ ), and  $R^2=40\%$  for communicating ( $F(1,69) = 45.71$ ,  $p < .001$ ). When all other variables were entered (Table 2), the inquiry pretest uniquely predicted 11%, 5%, and 8% of the variance for hypothesizing, interpreting data, and communicating, respectively.

Amongst the learner characteristics, self-reported academic efficacy, i.e., behaviors that would positively affect self performance and learning, is only a predictor for hypothesizing skill, and not for the other two inquiry skills. It added 4% explained variance in predicting hypothesizing skill ( $F(3,69) = 17.63, p < .001$ ) above the content and inquiry pretests.

Table 1: Predictors entered at each step in the forward linear regression when predicting each scientific inquiry skill.

Measure	Step	Predictor Added	$R^2$	$\Delta R^2$	F
Hypothesizing (N=70)	1	Content Pretest	.30	.30	29.40*
	2	Inquiry Pre	.41	.11	22.96*
	3	PALS4: Academic Efficacy	.45	.04	17.63*
Interpreting Data (N=70)	1	Content Pretest	.37	.37	40.37*
	2	Inquiry Pre	.42	.05	24.53*
Communicating Findings (N=70)	1	Content Pretest	.40	.40	45.71*
	2	Inquiry Pre	.48	.08	31.71*
Demonstrating CVS(N=63)	1	Content Pretest	.35	.35	32.67*
	2	PALS4: Skeptic School Relevance	.42	.08	22.09*

\* $p < .001$

Table 2: Forward regression for each inquiry skill predictor

Measure	Variable	B	SEB	$\beta$	Semipartial Corr <sup>2</sup>
Hypothesizing (N=70)	Content	1.95	.36	.55	.30*
	Pretest				
	Inquiry				
	Pre	.77	.22	.33	.11*
	PALS4:				
Interpreting Data (N=70)	Academic	.14	.07	.20	.04*
	Efficacy				
	Content	2.26	.36	.61	.37*
	Pretest				
Communicating Findings (N=70)	Inquiry	.55	.23	.24	.05*
	Pre				
	Content	2.19	.32	.63	.40*
	Pretest				
Demonstrating CVS(N=63)	Inquiry	.67	.20	.29	.08*
	Pre				
	Content	.16	.03	.59	.35*
	Pretest				
	PALS4	-.02	.01	-.28	.08*
	Skeptic				
	SchoolRel				

*Using Learner Characteristics to Predict Systematic Data Collection Behaviors.* We analyzed if any student learner characteristic subscales predicted authentic performance in our Archimedes activities using multiple regression. As mentioned previously, we used text replays to classify whether or not (1) students tested their hypotheses by targeting the independent variable as specified in their hypothesis, or (2) controlled for variables in each microworld activity. To form an estimate of students' proficiency at each skill, we used the average performance over all activities. For example, if we coded a student as controlling variables in 2 out of 3 activities, they would receive 67% for the proficiency estimate. In these analyses, we considered only students who completed all activities and learner characteristic surveys, leaving 63 students. Before performing regressions, we computed correlations between the learner characteristics and skill proficiency estimates. We found a very strong correlation between skill at designing controlled experiments ( $M = 0.57, SD = 0.39$ ) and skill at testing hypotheses (targeting IV) ( $M = 0.42, SD = 0.36$ ),  $r(63) = .83, p < .001$ ), indicating that students tend to possess both skills simultaneously. As a result, we consider only if learner characteristics predict skill at controlling for variables.

As expected, desirable learner characteristics, mastery orientation, academic efficacy and self-efficacy of scientific inquiry, were significantly and positively associated with skill at controlling for variables in our performance assessment. Correlations with mastery orientation ( $r = .29, p = .020$ ) and academic efficacy ( $r = .30, p = .019$ ) were slightly smaller but still significant. Self-efficacy at scientific inquiry correlated the strongest with skill at controlling for variables ( $r = .34, p = .007$ ). Additionally, lower values for less desirable self-reported learner characteristics like novelty avoidance ( $r = -.35, p = .005$ ), disruptive behavior ( $r = -.39, p = .002$ ), and being skeptical of school's relevance to success ( $r = -.39, p = .002$ ) were associated with greater skill at controlling variables. These significant correlations between behaviors and learner characteristics were low. Self-reporting as being performance-approach or performance-avoid oriented did not significantly correlate with authentic performance,  $p > .05$ . Finally, the content pretest ( $M = 6.22, SD = 1.43$ ) had a significant moderate correlation with this inquiry skill, ( $r = .59, p < .001$ ), indicating that students who knew the content material also may be better equipped to perform inquiry in this domain.

To determine which factors best predicted controlling for variables skill in our performance assessment, we performed a forward-selection within a linear regression framework. In forward multiple regression, at each stage the predictor that most increases explained variance ( $R^2$ ) with respect to the other variables already entered is added into the model. This value provided a benchmark for determining the relative contribution of predictors in explaining variance of authentic skill. A second benchmark, the square of the semi-partial correlation, was also used to examine relative

importance. After all variables are entered into the model, the square of the semi-partial correlation for a predictor gives the unique percentage of variance explained by that predictor, factoring out the shared variance with the other predictors.

As shown in Table 1, there are two significant predictors of each behavior. Most notably, the content pretest appeared as a significant variable predicting  $R^2=35\%$ , ( $F(1,61) = 32.67, p < .001$ ). Controlling for domain knowledge, being skeptical of school's relevance provided an additional  $\Delta R^2 = 8\%$  ( $F(2,60) = 22.09, p < .001$ ). As shown in Table 2, when all variables are entered, the content pretest uniquely predicted 35% of the variance in controlling for variables, and being skeptical of school's relevance uniquely predicted 8% of the variance. Thus, this learner characteristic can provide an additional, unique contribution towards predicting students' authentic inquiry performance above and beyond a baseline pretest score.

## Discussion

We addressed the influence of learner characteristics, namely learning orientation and self-efficacy on inquiry behaviors within a science learning environment. These learner characteristics were chosen because of their purported influence on students' cognitive processes during learning, including computer-based activities (Sins et al., 2007). Our goal in conducting this research was to collect data in order to begin to model these relationships to inform design and adaptive scaffolding for a wide variety of learners.

We used a combination of log files and open-text responses from our learning environment in order to:

- 1) conduct a fine-grained analysis of four scientific inquiry skills, namely *hypothesizing, interpreting data, communicating findings, and conducting controlled scientific experiments*. The latter involves two sub-skills namely, targeting the correct independent variable and using the control for variables strategy (CVS), and
- 2) test the relationships between these inquiry skills, learner characteristics, and content knowledge.

We found that our content pretest significantly predicted hypothesizing, data interpretation, communicating, and CVS better than our inquiry pretests. In terms of learner characteristics, one subscale, namely *academic efficacy*, made a unique contribution toward positively predicting students' skills at generating hypotheses beyond the content pretest. In addition, another subscale, namely, *skeptical of school relevance* made a unique contribution toward negatively predicting students' skills at CVS, thus, those who scored *higher* on skeptical about school's relevance scored *lower* on CVS. Taken at face value, this finding may suggest that those who are skeptical about school's relevance do not know the CVS skill. Another plausible interpretation is that these students did not engage in monitoring their inquiry and thus made careless errors in conducting their scientific trials. For example, a similar finding yielded data that suggested that mastery-oriented

students (those with the goal of deep learning) may not engage deeply with certain learning tasks unless they perceive them as useful for developing rich understanding (Crippen et al, 2009). In another study by our group, which addressed learning orientation and carelessness within microworlds, we developed detectors of students' carelessness at using CVS (Hershkovitz, Wixon, Baker, Gobert, & Sao Pedro, submitted). Here, carelessness is defined as it is in the cognitive tutors community, i.e., a behavior is deemed careless if the student had demonstrated poor performance on a skill for which they had shown mastery earlier. It is possible that this detector, when applied to the current data set, might elucidate the findings from the present study. Additional analyses are necessary to address this question.

As previously mentioned, it is our long-term goal to use various types of data in order to best support a wide range of students during inquiry. Prior literature suggests that students need support with monitoring their inquiry (de Jong, 2006), testing their stated hypotheses, and using the control for variables strategy (Sao Pedro et al., 2010). We (Gobert & Baker, 2010), as well as others (Crippen et al., 2009; Sins et al., 2007) believe that data regarding learner characteristics may be useful in order to inform both the design of instructional materials within our learning environment, as well as to adaptively support learners with specific learning orientations (Crippen et al., 2009). Moving forward, two key issues need to be addressed. The first is re-examining the face validity of self-report measures of learner characteristics such as learning orientation and self-efficacy. The second is to examine whether these learner characteristics play out differently in the context of science learning environments than they do in more traditional school tasks upon which many of the earlier studies on learner characteristics are based. As these issues are addressed, we can then begin to unpack the relationship between learner characteristics and fine-grained inquiry processes within science learning environments, and make a significant advance towards individualizing instruction for a broad range of learners.

## Acknowledgments

The Science Assistments project, lead by Janice Gobert, is generously funded by NSF-DRL#0733286, NSF-DGE# 0742503, NSF-DRL#1008649, and U.S. Dept of Ed. #R305A090170. Opinions expressed are those of the authors and do not necessarily reflect those of the agency.

## References

Baker, R. S. J. d., Corbett, A. T., Wagner, A. Z. (2006). Human Classification of Low-Fidelity Replays of Student Actions. *Proceedings of the Educational Data Mining Workshop at the 8th International Conference on Intelligent Tutoring Systems*, 29-36.

Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. New York, NY: W.H. Freeman and Company.

Crippen, K. J., Biesinger, K. D., Muis, K. R. & Orgill,

M. (2009). The role of goal orientation and self-efficacy in learning from web-based worked examples. *Journal of Interactive Learning Research* 20(4), 385–403.

de Jong, T. (2006). Computer simulations - Technological advances in inquiry learning. *Science*, 312, 532-533.

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.

Gobert, J., Heffernan, N., Baker, R., & Ruiz, C. (2007). *AMI: ASSISTments Meets Inquiry* (NSF-DRL #0733286). Proposal funded by National Science Foundation.

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

Gobert, J. & Baker, R. (2010). *Using Automated Detectors to Examine the Relationships Between Learner Attributes and Behaviors During Inquiry in Science* (NSF DRL-1008649). Proposal funded by the National Science Foundation.

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

Hershkovitz, A., Wixon, M., Baker, R. S. J. d., Gobert, J., & Sao Pedro, M. (Submitted, 2011). *Carelessness and goal orientation in a science microworld*. Submitted to The 15th International Conference on Artificial Intelligence in Education, Christchurch, New Zealand, June 27-July 1, 2011.

Ketelhut, D. J. (2007). The Impact of Student Self-efficacy on Scientific Inquiry Skills: An Exploratory Investigation in River City, a Multi-user Virtual Environment. *Journal of Science Education and Technology*, 16(1), 99-111.

McElhaney, K., & Linn, M. (2010). Helping Students Make Controlled Experiments More Informative. In K. Gomez, L. Lyons, & J. Radinsky (Ed.), *Learning in the Disciplines: Proceedings of the 9th International Conference of the Learning Sciences (ICLS 2010) - Volume 1, Full Papers* (pp. 786-793). Chicago, IL: International Society of the Learning Sciences.

Midgley, C., Maehr, M., Hruda, L., Anderman, E., Anderman, L., Freeman, K., et al. (2000). *Manual for the Patterns of Adaptive Learning Scales (PALS)*. Ann Arbor, MI: University of Michigan.

Montalvo, O., Baker, R. S. J. d., Sao Pedro, M. A., Nakama, A., Gobert, J. D. (2010). Identifying Student' Inquiry Planning Using Machine Learning. *Proceedings of the 3rd International Conference on Educational Data Mining*, pp. 141-150.

Nelson, B., & Ketelhut, D. (2008). Exploring embedded guidance and self-efficacy in educational multi-user virtual environments. *International Journal of Computer Supported Collaborative Learning*, 3(4), pp.413-427.

Payne, S. C., Youngcourt, S. S., & Beaubien, J. M. (2007). A meta-analytic examination of the goal orientation nomological net. *Journal of Applied Psychology*, 92(1), 128-150.

Perkins, D. (1986). *Knowledge as Design*. Hillsdale, NJ: Erlbaum.

Pintrich, P. R. (1999). The Role of Motivation in Promoting and Sustaining Self-Regulated Learning. *International Journal of Educational Research*, 31, 459-470.

Pintrich, P. R. (2000). Multiple Goals, Multiple Pathways: The Role of Goal Orientation in Learning and Achievement. *Journal of Educational Psychology*, 92, 544–555.

Razzaq, L., Feng, M., Nuzzo-Jones, G., Heffernan, N. T., Koedinger, K. R., et al. (2005). *The Assistment Project: Blending Assessment and Assisting*. In C.K. Looi, G. McCalla, B. Bredeweg, and J. Breuker (Eds.) *Proceedings of the 12th Artificial Intelligence In Education*, Amsterdam: ISO Press, pp.555-562.

Sao Pedro, M. A., Gobert, J. D., & Raziuddin, J. (2010). Comparing Pedagogical Approaches for the Acquisition and Long-Term Robustness of the Control of Variables Strategy. In K. Gomez, L. Lyons, & J. Radinsky (Ed.), *Learning in the Disciplines: Proceedings of the 9th International Conference of the Learning Sciences (ICLS 2010) - Volume 1, Full Papers* (pp. 1024-1031). Chicago, IL: International Society of the Learning Sciences.

Sao Pedro, M. A., Baker, R. S. J. d., Montalvo, O., Nakama, A., Gobert, J. D. (2010). Using Text Replay Tagging to Produce Detectors of Systematic Experimentation Behavior Patterns. *Proceedings of the 3<sup>rd</sup> International Conference on Educational Data Mining*, pp. 181-190.

Sins, P. H. M., van Joolingen, W. R., Savelbergh, E. R., van Hout-Wolters, B. (2007). Motivation and performance within a collaborative computer-based modeling task: Relations between students' achievement goal orientation, self-efficacy, cognitive processing, and achievement. *Contemporary Educational Psychology*, 33, 58-77.

Skaalvik, E. M. (1997). Self-Enhancing and Self-Defeating Ego Orientation: Relations With Task and Avoidance Orientation, Achievement, Self-Perceptions, and Anxiety, *Journal of Educational Psychology*, 89, 71–81.

Strand-Cary, M., & Klahr, D. (2008). Developing elementary science skills; Instructional effectiveness and path independence. *Cognitive Development* 23, 488-511.

Urdan, T. (2004). Predictors of Academic Self-Handicapping and Achievement: Examining Achievement Goals, Classroom Goal Structures, and Culture. *Journal of Educational Psychology*, 96, 251–264.

Wolters. C. (2004). Advancing Achievement Goal Theory: Using Goal Structures and Goal Orientations to Predict Students' Motivation, Cognition, and Achievement. *Journal of Educational Psychology*, 96, 236–250.