Goal Orientation, Self-Efficacy, and "Online Measures" in Intelligent Tutoring Systems

Stephen E. Fancsali[1] (sfancsali@carnegielearning.com)
Matthew L. Bernacki[2] (matt.bernacki@unlv.edu)
Timothy J. Nokes-Malach[3] (nokes@pitt.edu)
Michael Yudelson[1] (myudelson@carnegielearning.com)
Steven Ritter[1] (sritter@carnegielearning.com)
[1] Carnegie Learning, Inc., 437 Grant Street, Suite 918; Pittsburgh, PA 15219 USA
[2] Department of Educational Psychology & Higher Education, University of Nevada, Las Vegas; Las Vegas, NV 89154 USA
[3] Department of Psychology and Learning Research and Development Center, University of Pittsburgh; Pittsburgh, PA 15260 USA

Abstract

While goal orientation and related factors like learner self-efficacy are of great interest to learning science researchers, some voice concerns regarding the measurement of such factors using self-report questionnaires. To address these concerns, recent work has explored the use of behavioral indicators like hint-seeking and glossary use in intelligent tutoring systems like Carnegie Learning's Cognitive Tutor® (CT) as alternative, "online" measures of goal orientation. We re-examined this approach by measuring 273 CT users' achievement goals and self-efficacy judgments via embedded questionnaires and their hint-seeking and glossary use via log data. Using graphical causal models and linear structural equation models to observe structural relationships among goal orientations, self-efficacy, behaviors, and learning outcomes, we found that tracing orientations via "online measures" is more nuanced than perhaps previously appreciated. We describe complex relations observed in the model among motivations, behaviors, and outcomes and discuss the implications for the online measurement of motivation.

Keywords: Goal Orientation; Motivation; Self-Efficacy; Non-Cognitive Factors; Intelligent Tutoring Systems; Structural Equation Models; Graphical Causal Models.

Introduction

One well-studied aspect of motivation for learning focuses on individuals' achievement goals when approaching a learning task. Dweck (1986) provides a distinction between mastery and performance goal orientations. Learners have a mastery goal orientation when they seek to understand (i.e., master) a particular task or domain of interest. Those who seek to perform better relative to others have a performance goal orientation. Later work added another dimension of variation: a "valence" of either approaching success or avoiding failure (Elliot & McGregor, 2001). Learner goals corresponding to a mastery approach are those aimed at *developing* competence with respect to a task or learning objective, perhaps over a previous personal level of competence or other self-imposed criterion for task-mastery (Ames, 1992; Elliot, 1999); performance approach goals seek to *demonstrate* competence by outperforming peers. Learners who endorse performance avoidance goals seek to demonstrate that they are not any less competent than peers.

Self-report questionnaires are commonly used to measure goal orientation. Generally, questionnaires are provided to learners either before or after a learning task. However, goal orientation can change dynamically as learners progress through a learning experience and have been shown to vary over longer time periods (e.g., a semester; Richardson, 2004; Fryer & Elliot, 2007; Muis & Edwards, 2009).

Consequently, recent work (Otieno, Schwonke, Salden, & Renkl, 2013) suggests that, given changing or state-like aspects of goal orientations, fine-grained, "online" measures of goal orientation (i.e., those extracted from software log "traces") may be a fruitful supplement to, and a potentially better measure than, questionnaire data in learning environments like intelligent tutoring systems (ITSs). While we agree that developing and validating appropriate "online" measures of goal orientation as well as other motivational, metacognitive, and cognitive processes is an important line of research, we suggest that relatively simple, proposed online measures may not provide a sufficiently nuanced assessment of underlying phenomena and may conflate a motivational construct with a behavior resulting from one or more motivations.

We considered data from a study conducted by the second and third authors that addresses state-like aspects of goal orientation using online, in-tutor (i.e., between units of mathematics content) questionnaires in Carnegie Learning's Cognitive Tutor® (CT) (Carnegie Learning, 2012; Ritter, Anderson, Koedinger, & Corbett, 2007) ITS for mathematics. We adopted a path analytic approach using structural equation models to investigate relationships among a variety of self-reports of students' motivation, online measures of students' behavior in and interaction with the CT, and performance outcomes. We specified a structural equation model by learning a set of qualitative causal structures consistent with both data and background knowledge using the framework of semi-automated, algorithmic search for graphical causal models (Spirtes, Glymour, & Scheines 2000; Pearl, 2009).

We evaluate the proposal of Otieno, et al. (2013) that hint and glossary use in the CT may serve as online indicators of student motivation (i.e., goal orientation) and found that their proposed mapping of traced behavior to motivational construct could not be reproduced with these data. Glossary use was possibly an effect (and thereby a possible indicator) of learner self-efficacy, but not goal orientation. Similarly, hint-use may serve as a weak indicator of self-efficacy, but our analysis of correlations and structural relationships among measures suggest that the use of glossary use and hint use as behavioral indicators of specific learner motivations (i.e., mastery-approach and performance-approach goals) may not be appropriate. We conclude by suggesting several important problems to be addressed by future research that aims to develop online measures of motivational, affective, cognitive and metacognitive processes.

Background and Motivation

Two primary goals motivated the study described in the following section: (1) to address aforementioned shortcomings of self-report questionnaires by measuring goal orientation and self-efficacy at both a finer level of granularity and over a longer period of time than in previous studies and (2) to determine associations and relationships among these factors (measured at different levels of granularity) and several learning outcomes and ITS behavioral variables.

In prior work, Bernacki, Nokes-Malach, & Aleven (2013) found that when achievement goals are reported with different levels of specificity (i.e., achievement goals for mathematics versus achievement goals for a CT unit), the strength of association between achievement goals and behaviors differs. This suggests that different levels of self-report can serve as useful predictors of learning behavior. Additionally, a second study that examined the stability and change in achievement goals over CT units (Bernacki, Nokes-Malach, & Aleven, in press) using reliable change indices (Caspi, Roberts, & Shiner, 2005) revealed that the majority of students' achievement goals change reliably from one unit to the next. These findings suggest that the achievement goals individuals report may be determined not only by the specificity of the reporting criteria, but also by the features of the task (i.e., ITS units). In this study, we examine how achievement goals and self-efficacy for math and for CT math units predict students' behaviors in those units, including hint and glossary use, and how motivations and behaviors predict performance.

Further, rich data collected in the study also allows us to address recent questions about online measures of motivational factors, including goal orientation. To avoid the shortcomings associated with reliance on self-report questionnaires to assess factors like goal orientation, Otieno, et al. (2013) and Zhou & Winne (2012) have suggested that online traces (e.g., measured indicators from software log files) may provide better, less obtrusive means by which to assess student motivation. Specifically, Otieno and colleagues argue that hint use in the CT may serve as an indicator of performance goal orientation and glossary use as an indicator of mastery goal orientation.

While we agree that online traces are advantageous and an important topic for future research, we believe that the frequency of hint use is, at best, both too coarse and too "noisy" to be an indicator of a performance goal orientation. When behaviors are used as a trace of a motivational construct, it is often the case that the behavior traced is a theorized *product* of a motivational state and not necessarily a characteristic of one who experiences the state. This conflation of motivational state and resulting behavior is akin to identifying an illness by a single symptom of such an illness, ignoring that many other illnesses may produce the same single symptom.

Learners' decision to use hints may stem from a variety of motivations. Those who seek to improve their performance in a CT unit may abuse hints (Aleven & Koedinger, 2000), in an attempt to "game the system" (Baker, Corbett, Koedinger, & Wagner, 2004) and increase estimates of their skillfulness without actually trying to learn targeted skills. Such behavior might reflect a performance orientation and the absence of a mastery orientation if the student abused hints because they perceived it as means to achieve progress relative to their peers. This could be perceived as evidence of a "shallow process" which has been associated with performance goals (Elliot, 1999).

However, Otieno, et al. (2013) also posit that hint use is a better indicator of *performance approach* rather than *mastery approach* because students do not often reflect upon hints presented to them. Recent research indicates this is not always the case. A response time model developed by Shih, Koedinger, and Scheines (2011) demonstrates instances when learners likely reflect on the hints they request, especially after "bottom-out" hints, which provide students with the correct answer to a step of a problem. They propose this behavior may be evidence that a student has adopted a strategy of seeking worked examples, a "deep processing" strategy to improve their understanding (i.e., associated with mastery approach goals; Elliot, 1999). In sum, theory and recent empirical evidence suggest the relations between achievement goals and hint behaviors are complex. Depending on the way hints are used, hint use could serve as an indication of very different achievement goals. There is less work examining relationships between glossary use in CT or similar features in ITSs and motivation, so we further explore possible relationships here.

Study

In light of the potential for complex relations between achievement goals, other motivational processes, and learning behavior, we adopted an exploratory approach to observing learners' achievement goals and learning behavior in CT units and examined the relations between motivations, behaviors and outcomes. We next describe the sample, CT learning environment, self-report questionnaires, achievement data, and our analytical approach.

Participants

Our sample consisted of 273 middle and high school students taking pre-algebra, algebra, and geometry courses that use the CT regularly (i.e., two class periods per week) at a suburban high school in western Pennsylvania. The student population was primarily Caucasian (97%) and included an approximately balanced gender ratio.

Cognitive Tutor (CT)

CT is an ITS for mathematics with hundreds of thousands of middle and high school users throughout the United States. The CT provides adaptive tutoring by tracking mastery of individual knowledge components (KCs) or skills as learners progress through mathematics content, using a probabilistic framework called Bayesian Knowledge Tracing (Corbett & Anderson 1995). Mastery is operationalized as a learner reaching an assessed 0.95 probability of KC knowledge.

At a higher level, each mathematics sub-discipline (e.g., algebra) is divided into units, and units are divided into (roughly topical) sections composed of many problems. Each problem in the CT has KCs associated with it, so performance on opportunities to practice KCs is tracked as learners solve particular steps of problems (e.g., a cell in the table in the screenshot in Figure 1). The CT provides immediate feedback about correctness at each step of a problem (all incorrect responses counting as errors), and context-sensitive hints are available for each step of a problem a learner attempts to solve. In some cases, immediate, "just-in-time" feedback is also provided to learners when particular errors are made. A learner must be judged by the CT to have achieved mastery of all KCs associated with a particular section before "graduating" to the following section. Having graduated from all sections in a unit, the learner graduates to the following unit.

Self-Report Questionnaire Method

In the original study, middle and high school learners completed a series of self-report questionnaires within the CT assessing achievement goals and self-efficacy over the course of several units of instruction in the CT (Bernacki, Nokes-Malach, & Aleven, 2013). Learners responded to "domain-level" items assessing their goal orientation and self-efficacy (i.e., with respect to mathematics) as well as "unit-level" questionnaire items (i.e., about the particular CT unit). Unit-level questionnaire content alternated between a report on achievement goals in one unit and an assessment of self-efficacy in the next.

Achievement Goals Students' achievement goals were assessed using original items for mastery approach, performance approach, and performance avoidance subscales of the Achievement Goals Questionnaire – Revised (Elliot & Murayama, 2008). Students responded by rating their agreement with items on an integer scale of 1 (not at all true of me) to 7 (very true of me). The domain-level subscales from this questionnaire included items with "this [mathematics] class" as the referent:

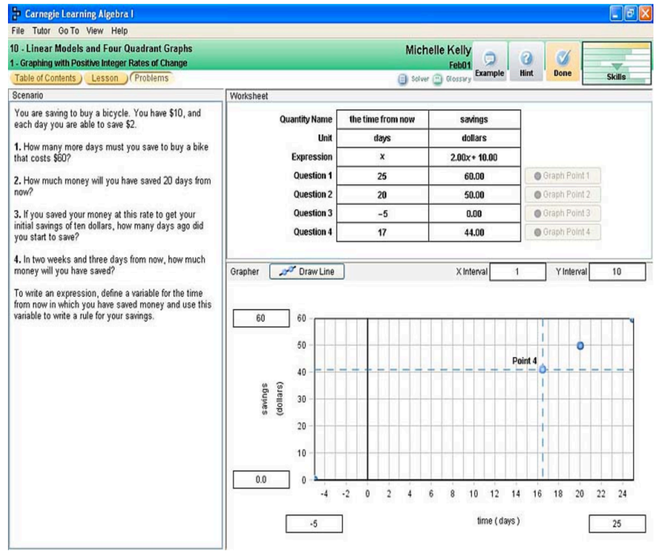

Figure 1: Problem solving screenshot of CT Algebra

- My aim is to completely master the material presented in this class. (mastery approach)
- I am striving to do well compared to the other students. (performance approach)
- My goal is to avoid performing poorly compared to others. (performance avoidance)

For the unit-level questionnaires, the terms "class/course" were replaced by the term "unit" for items in each subscale.

Self-Efficacy To measure learner self-efficacy, the second author designed survey items, according to recommendations provided by Bandura (2006), specifically for the mathematics domain and ITS units of mathematics instruction. Students responded in accordance with their agreement with items in the form of integer ratings between 1 (not at all true) and 9 (completely true). The domain-level items were worded with respect to the math course (e.g., mastery-approach item: "I am confident that I will do well in math class."). For unit-level items, appropriate changes were made to the domain-level items (e.g., "I am confident that I will do well on units like this one.").

Data

We considered domain- and unit-level measures of goal orientation and self-efficacy, log data of student interaction with the CT, prior-year math course grades, and final course grades for the math course in which the CT was used.

From CT log data, we extracted the number of errors made, hints requested, and problems required to finish each unit, as well as the count of the number of times that students read glossary entries. Hints, errors, and problems were normalized over units because of the differing numbers of problems required to complete different units due to factors such as unit content and mode of delivery (e.g., equation solving vs. word problems, etc.). We take the average of normalized student scores over all units to produce a single variable representing each type of tutor action. Individual unit-level questionnaire items were summed and normalized per construct, per unit, as well, to produce a single score per construct across all units. Three variables were normalized over all students: glossary use, prior final mathematics grade, and course final grade. The normalization of each variable also provides for better interpretability of estimated parameters in the statistical model we present in the following section.

Causal Graphs & Path Analytic Approach

We adopted a path analysis approach using linear structural equation models that allowed us to investigate a variety of questions, including those about mediation, about features of interest. Such an approach has been adopted in a variety of experimental and observational studies and fruitfully used to analyze log data from ITSs like the CT (e.g., Rau & Scheines, 2012; Rau, Scheines, Aleven, & Rummel, 2013).

Lacking a strong theory about specific causal links among features of interest and/or mediation relationships among them (and the mixed bag of prior results) to fully specify the a structural equation model, we adopted a data-driven approach to search for qualitative causal structure(s), represented by graphical causal models, consistent with data and our available background knowledge (Spirtes, et al., 2000; Pearl 2009). Qualitative causal structure of a linear structural equation model can be represented by a directed acyclic[1] graph (DAG); under the causal interpretation of a DAG, directed edges represent direct causal links relative to the set of variables or features in the DAG.

Assuming multivariate normal distributions and linear causal dependencies, asymptotically reliable search procedures (Spirtes, et al., 2000) are available to infer the equivalence class of DAGs consistent with observed (conditional) independence relationships and available background knowledge (e.g., time-ordering). The equivalence class, called a pattern (Spirtes, et al., 2000), provides the set of DAGs that are observationally indistinguishable (i.e., that entail the same set of (conditional) independence constraints).

We used GES[2] (Chickering, 2002), an algorithm that finds the pattern that optimizes the Bayesian Information Criterion (BIC) (Schwarz, 1978) score. Beyond parametric assumptions that causal dependencies are linear and of multivariate normal distributions, one caveat is that it is assumed that there are no unmeasured common causes of measured variables. Since this latter assumption almost certainty does not hold in this domain, we later discuss relaxing this assumption.

[1] i.e., no "feedback loops"

[2] available in the Tetrad IV software & suite of algorithms (http://www.phil.cmu.edu/projects/tetrad)

Results

Since all DAGs in the pattern learned by GES will fit the data equally well, we arbitrarily orient those edges left un-oriented by the algorithm; in this case, only those edges between the measures of mastery approach, performance approach, and performance avoidance (i.e., all included measures of goal orientation) are un-oriented by GES.

We use the resulting DAG to specify a linear structural equation model; the estimated model (Figure 2) fits the data as assessed by a chi-square statistical test of whether there is a significant difference between the implied covariance matrix of the estimated linear model and the observed covariance matrix among measured variables [$\chi^2(43) = 49.19$, $p = .239$] (Bollen, 1989).

Regardless of whether particular edges are interpreted as direct causal links, the qualitative structure (i.e., conditional independencies implied by) and parameter estimates of the model in Figure 2 lead us to several conclusions about self-efficacy, goal orientation, and previously proposed online measures thereof. We then consider relaxing the assumption of "no unmeasured common causes" of measured variables and the "goal complex" associated with measures of goal orientation before discussing future research.

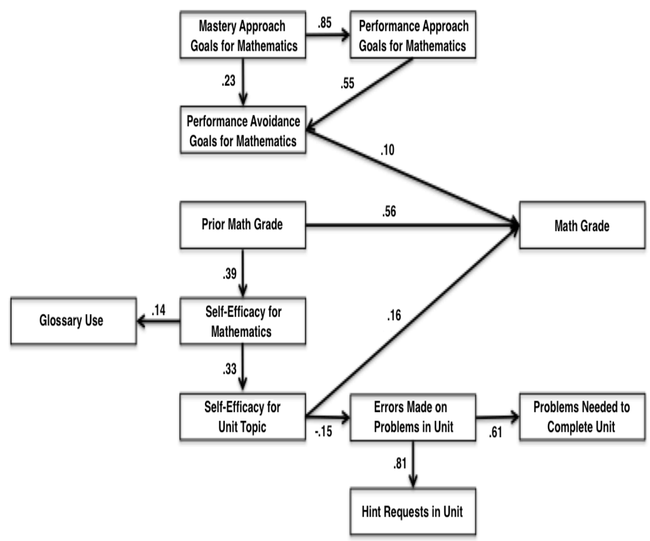

Figure 2: Illustration of estimated linear structural equation model

Goal Orientation, Self-Efficacy, and Learning

Our findings are consistent with prior work (Bernacki, et al., 2013) demonstrating that measures at different levels of granularity (here, domain-level versus unit-level self-efficacy) provide different information about students' in-tutor behavior. We found a direct link between the unit-level self-efficacy and learning (i.e., *Math Grade*). We describe the path to grades from a "goal complex" view below.

Hint-Seeking & Glossary Use as Online Measures

Domain-level self-efficacy was weakly associated with glossary use ($r = .15$, $p = .014$); conditional on domain-level self-efficacy, glossary use was independent of all other measured variables. In contrast to the argument of Otieno, et al. (2013), glossary use may be conceived of as an online measure of domain-level self-efficacy, but we found no evidence of a direct link to mastery goals. The weak correlation suggests that glossary use was, at best, a noisy measure of self-efficacy. Further, hint use was independent of all other variables conditional on our measure of errors. Errors were weakly linked to unit-level self-efficacy ($r = -.18$, $p = .002$); we found no significant correlations between hint use and goal orientation.

Unobserved Common Causes & Goal "Complex"

To consider the robustness of our search for qualitative causal structure, we also used a constraint-based search algorithm called FCI (Spirtes, et al., 2000) that allows for the possibility of unmeasured common causes among measured variables. The qualitative structure of the result of constraint-based search is similar to that of GES, but it suggests that we cannot tell from observational data alone whether possible causal links between *Math Grade* and prior knowledge (i.e., *Prior Math Grade*), unit-level self-efficacy, and a posited link between domain-level self-efficacy (not in the model from GES) are confounded by unmeasured common causes. This is unsurprising as we include measured proxies for latent phenomena, and other latent phenomena may be responsible for such correlations. FCI also omits any link between measures of goal orientation and *Math Grade*, but this may also be a product of our modeling of the underlying latent phenomena with such measured proxies. Further, FCI suggests there is better evidence that two links are not confounded: (1) between domain-level self-efficacy and glossary use and (2) between domain-level self-efficacy and unit-level self-efficacy. This bolsters the possibility of an online measure of domain-level self-efficacy, but does nothing to cure the weakness of this link.

Another theoretical and statistical complication is raised by past work that suggests students often endorse multiple goals, resulting in goal scores that are highly correlated and indicative of a "goal complex" (Barron & Harackiewicz, 2001; Senko, Hulleman, & Harackiewicz, 2011). Relatively large observed correlations among our measures of goal orientation, coupled with results of GES, FCI, and factor analytic techniques provide evidence for a goal orientation complex. Our results suggest that one or more latent variables could explain statistical dependencies among the three measures of goal orientation, rather than each serving as a proxy for a particular goal orientation/valence, but theoretical, statistical, and measurement questions remain.

Discussion

We agree with the assessment of Otieno, et al. (2013) that latent phenomena like motivation may be better measured by some combination of self-report questionnaires and online traces gleaned from rich ITS log data. Whereas Otieno and colleagues argue that hint and glossary usage are potential behavioral traces of performance approach and mastery approach goals, we find that when multiple measures of motivation were included in a model, the traced behaviors associated more strongly with self-efficacy than either students' achievement goals for math or ITS units.

We suggest that, in the context of ITSs for mathematics, attempting to trace motivation via these behaviors produces weak and noisy measures. This stands in contrast to work conducted by Zhou and Winne (2012) that traced learners' goal orientations in a hypertext reading task by explicitly labeled annotations. In the context of a reading comprehension task where learners could tag a passage as important for performance (e.g., "important to know for test") or for mastery (e.g., "I want to know more about this."), the process of aligning behavioral traces to features of learners' goal orientations is clearer. However, in the context of mathematics, researchers have not conducted a similar study where one's intentions for using a resource can be explicitly labeled. Separate buttons could be developed that provide a hint "to get this problem right" or "to understand this concept better," but at present, learners' (many) motivations for using a hint or glossary tool are unknowable.

Alternatively, we suggest that more sophisticated feature engineering (e.g., including features that capture timing of particular actions) may be used alongside self-report data to produce "sensor-free" detectors of motivational factors akin to processes used to detect when learners are "gaming the system" (e.g., Baker, et al., 2004). Such means may provide ways to "detect" student motivation and eliminate the need to conduct obtrusive, time-consuming questionnaires. In the future, path analytical approaches incorporating questionnaires and/or detectors could be used to improve our understanding of the implications that learners' motivations have for behaviors and learning outcomes.

Acknowledgments

The 2nd and 3rd authors were funded by a grant from the Pittsburgh Science of Learning Center (Grant #0836012). The 1st, 4th, and 5th authors were funded by the U.S. Army Research Laboratory's Advanced Distributed Learning Initiative Contract #W911QY-13-C-0026 under ADL BAA 12-001. We thank Vincent Aleven for his involvement in the design of the original study on which this paper builds.

References

Aleven, V., & Koedinger, K.R. (2000). Limitations of student control: do students know when they need help? *Proceedings of the 5th International Conference on Intelligent Tutoring Systems* (pp. 292-303). Berlin: Springer-Verlag.

Ames, C. (1992). Classrooms: goals, structures, and student motivation. *Journal of Educational Psychology, 84*, 261-271.

Baker, R.S., 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* (pp. 383-390). New York: ACM.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.) *Self-Efficacy Beliefs for Adolescents*. IAP.

Barron, K. E., & Harackiewicz, J. M. (2001). Achievement goals and optimal motivation: testing multiple goal models. *Journal of Personality and Social Psychology*, *80*, 706-722.

Bernacki, M.L., Nokes-Malach, T.J., & Aleven, V. (2013). Fine-grained assessment of motivation over long periods of learning with an intelligent tutoring system: methodology, advantages, and preliminary results. In R. Azevedo & V. Aleven (Eds.) *International Handbook of Metacognition and Learning Technologies*. New York: Springer.

Bernacki, M.L., Nokes-Malach, T.J., & Aleven, V. (in press). Stability and change in adolescents' task-specific achievement goals for learning mathematics with an intelligent tutoring system. *Computers in Human Behavior.* doi: 10.1016/2014.04.009

Bollen, K. (1989). *Structural Equations with Latent Variables*. Wiley.

Carnegie Learning (2012). Cognitive Tutor [software]. Pittsburgh, PA: Carnegie Learning, Inc.

Caspi, A., Roberts, B.W., & Shiner, R.L. (2005). Personality development: stability and change. *Annual Review of Psychology, 56*, 453-484.

Chickering, D.M. (2002). Optimal structure identification with greedy search. *Journal of Machine Learning Research, 3*, 507-554.

Corbett, A.T., & Anderson, J.R. (1995). Knowledge tracing: modeling the acquisition of procedural knowledge. *User Modeling and User Adapted Interaction, 4*, 253-278.

Dweck, C.S. (1986). Motivational processes affecting learning. *American Psychologist, 41*, 1040-1048.

Elliot, A.J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist, 34*, 169-189.

Elliot, A.J., & McGregor, H.A. (2001). A 2X2 achievement goal framework. *Journal of Personality and Social Psychology, 80*, 501-519.

Elliot, A.J., & Murayama, K. (2008). On the measurement of achievement goals: critique, illustration, and application. *Journal of Educational Psychology, 100*, 613-628.

Fryer, J.W., & Elliot, A.J. (2007). Stability and change in achievement goals. *Journal of Educational Psychology, 99*, 700-714.

Muis, K.R., & Edwards, O. (2009). Examining the stability of achievement goal orientation. *Contemporary Educational Psychology, 34*, 265-277.

Otieno, C., Schwonke, R., Salden, R., & Renkl, A. (2013). Can help seeking behavior in intelligent tutoring systems be used as an online measure for goal orientation? *Proceedings of the 35th Annual Meeting of the Cognitive Science Society* (pp. 1103-1108). Austin, TX: Cognitive Science Society.

Pearl, J. (2009). *Causality: Models, Reasoning, and Inference*. 2nd Edition. Cambridge: Cambridge UP.

Rau, M., & Scheines, R. (2012). Searching for Variables and Models to Investigate Mediators of Learning from Multiple Representations. In K. Yacef, O. Zaïane, A. Hershkovitz, M. Yudelson, & J. Stamper (Eds.) *Proceedings of the 5th International Conference on Educational Data Mining*. (pp. 110-117).

Rau, M., Scheines, R., Aleven, V., & Rummel, N. (2013). Does Representational Understanding Enhance Fluency Or Vice Versa? Searching for Mediation Models. In S. K. D'Mello, R. A. Calvo & A. Olney (Eds.) *Proceedings of the 6th International Conference on Educational Data Mining* (pp. 161-169). International Educational Data Mining Society.

Richardson, J.T.E. (2004). Methodological issues in questionnaire-based research on student learning in higher education. *Educational Psychology Review, 16*, 347-358.

Ritter, S., Anderson, J.R., Koedinger, K.R., & Corbett, A.T. (2007). Cognitive Tutor: applied research in mathematics education. *Psychonomic Bulletin and Review, 14*, 249-255.

Schwarz, G.E. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461-464.

Senko, C., Hulleman, C. S., & Harackiewicz, J. M. (2011). Achievement goal theory at the crossroads: old controversies, current challenges, and new directions. *Educational Psychologist*, *46*, 26-47.

Shih, B., Koedinger, K.R., & Scheines, R. (2011). A response-time model for bottom-out hints as worked examples. In C. Romero, S. Ventura, M. Pechenizkiy, & R.S.J.d. Baker (Eds.) *Handbook of Educational Data Mining.* Boca Raton, FL: CRC.

Spirtes, P., Glymour, C., & Scheines, R. (2000). *Causation, Prediction, and Search*. 2nd Edition. Cambridge, MA: MIT.

Zhou, M., & Winne, P.H. (2012). Modeling academic achievement by self-reported versus traced goal orientation. *Learning and Instruction*, *22*, 413-419.