How to Submit Proof Corrections Using Adobe Reader

Using Adobe Reader is the easiest way to submit your proposed amendments for your IGI Global proof. If you don't have Adobe Reader, you can download it for free at http://get.adobe.com/reader/. The comment functionality makes it simple for you, the contributor, to mark up the PDF. It also makes it simple for the IGI Global staff to understand exactly what you are requesting to ensure the most flawless end result possible.

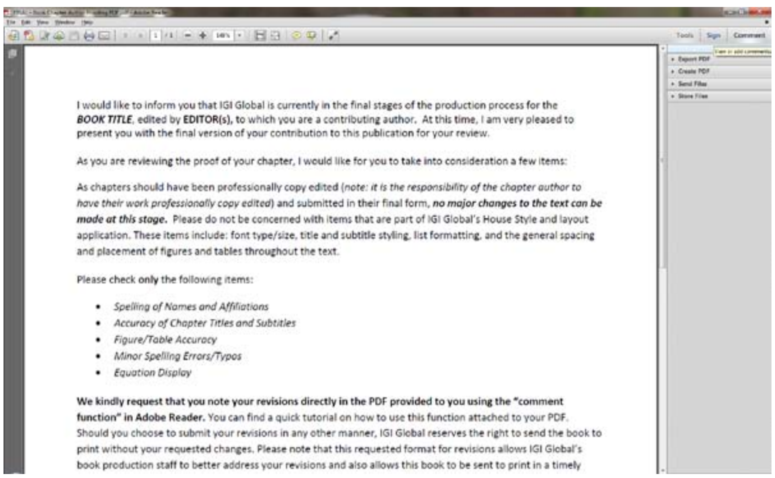
Please note, however, that at this point in the process the only things you should be checking for are:

Spelling of Names and Affiliations, Accuracy of Chapter Titles and Subtitles, Figure/Table Accuracy, Minor Spelling Errors/Typos, Equation Display

As chapters should have been professionally copy edited and submitted in their final form, please remember that ***no major changes to the text can be made at this stage*.**

Here is a quick step-by-step guide on using the comment functionality in Adobe Reader to submit your changes.

1. Select the **Comment** bar at the top of page to View or Add Comments. This will open the **Annotations** toolbar.



I would like to inform you that IGI Global is currently in the final stages of the production process for the ***BOOK TITLE***, edited by **EDITOR(s)**, to which you are a contributing author. At this time, I am very pleased to present you with the final version of your contribution to this publication for your review.

As you are reviewing the proof of your chapter, I would like for you to take into consideration a few items:

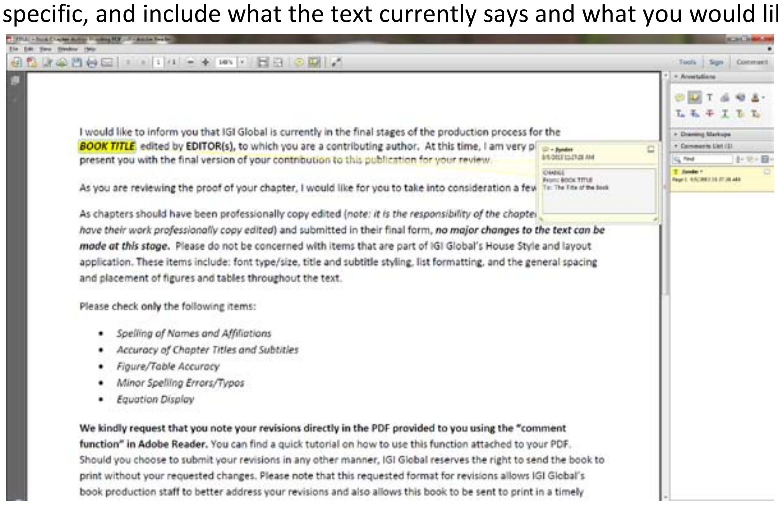
As chapters should have been professionally copy edited (*note: it is the responsibility of the chapter author to have their work professionally copy edited*) and submitted in their final form, ***no major changes to the text can be made at this stage***. Please do not be concerned with items that are part of IGI Global's House Style and layout application. These items include: font type/size, title and subtitle styling, list formatting, and the general spacing and placement of figures and tables throughout the text.

Please check only the following items:

- *Spelling of Names and Affiliations*
- *Accuracy of Chapter Titles and Subtitles*
- *Figure/Table Accuracy*
- *Minor Spelling Errors/Typos*
- *Equation Display*

We kindly request that you note your revisions directly in the PDF provided to you using the "comment function" in Adobe Reader. You can find a quick tutorial on how to use this function attached to your PDF. Should you choose to submit your revisions in any other manner, IGI Global reserves the right to send the book to print without your requested changes. Please note that this requested format for revisions allows IGI Global's book production staff to better address your revisions and also allows this book to be sent to print in a timely

2. To note text that needs to be altered, like a subtitle or your affiliation, you may use the **Highlight Text** tool. Once the text is highlighted, right-click on the highlighted text and add your comment. Please be specific, and include what the text currently says and what you would like it to be changed to.



I would like to inform you that IGI Global is currently in the final stages of the production process for the ***BOOK TITLE***, edited by **EDITOR(s)**, to which you are a contributing author. At this time, I am very p
present you with the final version of your contribution to this publication for your review.

As you are reviewing the proof of your chapter, I would like for you to take into consideration a few

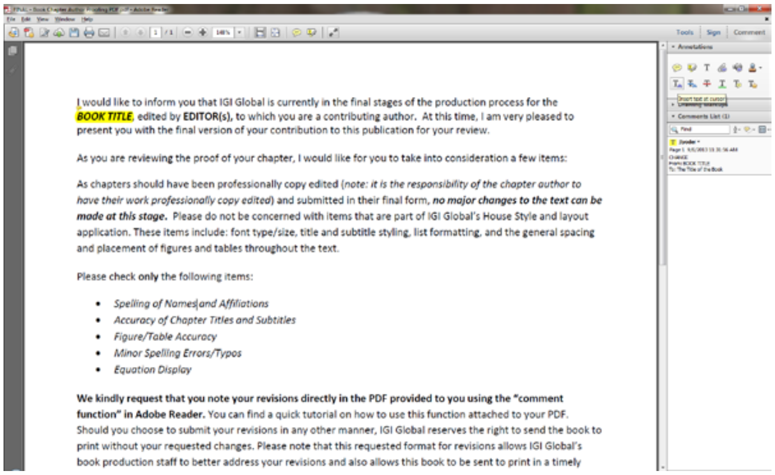
As chapters should have been professionally copy edited (*note: it is the responsibility of the chapte*
have their work professionally copy edited) and submitted in their final form, ***no major changes to the text can be made at this stage***. Please do not be concerned with items that are part of IGI Global's House Style and layout application. These items include: font type/size, title and subtitle styling, list formatting, and the general spacing and placement of figures and tables throughout the text.

Please check only the following items:

- *Spelling of Names and Affiliations*
- *Accuracy of Chapter Titles and Subtitles*
- *Figure/Table Accuracy*
- *Minor Spelling Errors/Typos*
- *Equation Display*

We kindly request that you note your revisions directly in the PDF provided to you using the "comment function" in Adobe Reader. You can find a quick tutorial on how to use this function attached to your PDF. Should you choose to submit your revisions in any other manner, IGI Global reserves the right to send the book to print without your requested changes. Please note that this requested format for revisions allows IGI Global's book production staff to better address your revisions and also allows this book to be sent to print in a timely

3. If you would like text inserted, like a missing coma or punctuation mark, please use the **Insert Text at Cursor** tool. Please make sure to include exactly what you want inserted in the comment box.


I would like to inform you that IGI Global is currently in the final stages of the production process for the ***BOOK TITLE***, edited by **EDITOR(s)**, to which you are a contributing author. At this time, I am very pleased to present you with the final version of your contribution to this publication for your review.

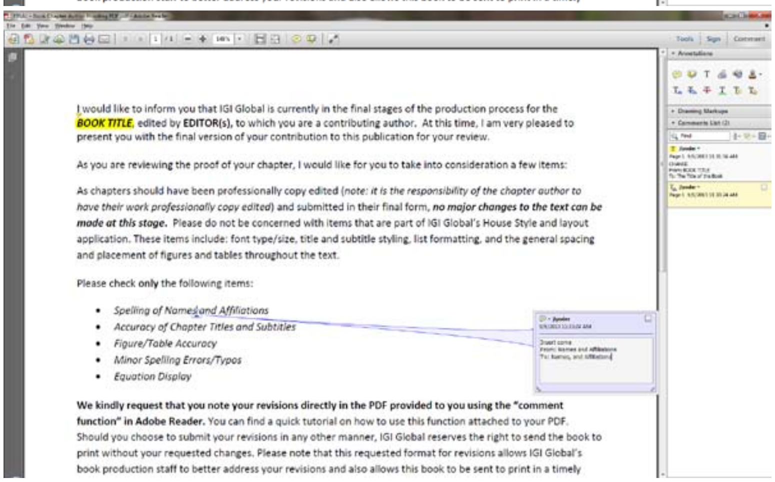
As you are reviewing the proof of your chapter, I would like for you to take into consideration a few items:

As chapters should have been professionally copy edited (*note: it is the responsibility of the chapter author to have their work professionally copy edited*) and submitted in their final form, ***no major changes to the text can be made at this stage.*** Please do not be concerned with items that are part of IGI Global's House Style and layout application. These items include: font type/size, title and subtitle styling, list formatting, and the general spacing and placement of figures and tables throughout the text.

Please check only the following items:

- *Spelling of Names and Affiliations*
- *Accuracy of Chapter Titles and Subtitles*
- *Figure/Table Accuracy*
- *Minor Spelling Errors/Typos*
- *Equation Display*

We kindly request that you note your revisions directly in the PDF provided to you using the "comment function" in Adobe Reader. You can find a quick tutorial on how to use this function attached to your PDF. Should you choose to submit your revisions in any other manner, IGI Global reserves the right to send the book to print without your requested changes. Please note that this requested format for revisions allows IGI Global's book production staff to better address your revisions and also allows this book to be sent to print in a timely


I would like to inform you that IGI Global is currently in the final stages of the production process for the ***BOOK TITLE***, edited by **EDITOR(s)**, to which you are a contributing author. At this time, I am very pleased to present you with the final version of your contribution to this publication for your review.

As you are reviewing the proof of your chapter, I would like for you to take into consideration a few items:

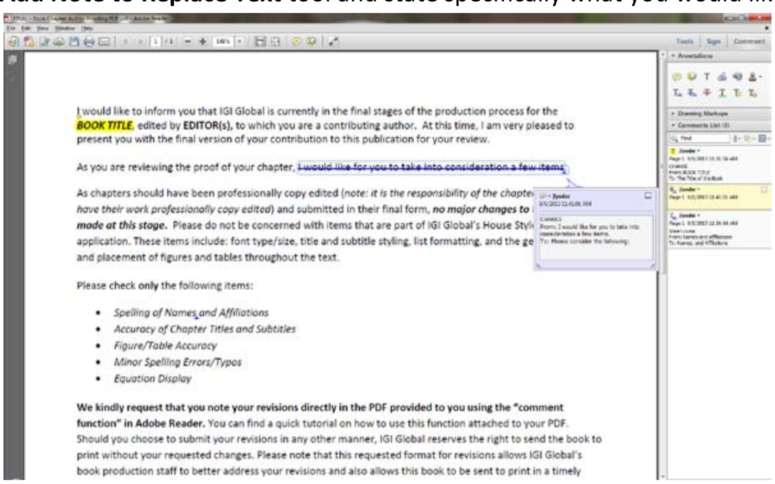
As chapters should have been professionally copy edited (*note: it is the responsibility of the chapter author to have their work professionally copy edited*) and submitted in their final form, ***no major changes to the text can be made at this stage.*** Please do not be concerned with items that are part of IGI Global's House Style and layout application. These items include: font type/size, title and subtitle styling, list formatting, and the general spacing and placement of figures and tables throughout the text.

Please check only the following items:

- *Spelling of Names and Affiliations*
- *Accuracy of Chapter Titles and Subtitles*
- *Figure/Table Accuracy*
- *Minor Spelling Errors/Typos*
- *Equation Display*

We kindly request that you note your revisions directly in the PDF provided to you using the "comment function" in Adobe Reader. You can find a quick tutorial on how to use this function attached to your PDF. Should you choose to submit your revisions in any other manner, IGI Global reserves the right to send the book to print without your requested changes. Please note that this requested format for revisions allows IGI Global's book production staff to better address your revisions and also allows this book to be sent to print in a timely

4. If you would like text removed, such as an erroneous duplicate word or punctuation mark, please use the **Add Note to Replace Text** tool and state specifically what you would like removed.


I would like to inform you that IGI Global is currently in the final stages of the production process for the ***BOOK TITLE***, edited by **EDITOR(s)**, to which you are a contributing author. At this time, I am very pleased to present you with the final version of your contribution to this publication for your review.

As you are reviewing the proof of your chapter, ~~I would like for you to take into consideration a few items:~~

As chapters should have been professionally copy edited (*note: it is the responsibility of the chapter author to have their work professionally copy edited*) and submitted in their final form, ***no major changes to the text can be made at this stage.*** Please do not be concerned with items that are part of IGI Global's House Style and layout application. These items include: font type/size, title and subtitle styling, list formatting, and the general spacing and placement of figures and tables throughout the text.

Please check only the following items:

- *Spelling of Names and Affiliations*
- *Accuracy of Chapter Titles and Subtitles*
- *Figure/Table Accuracy*
- *Minor Spelling Errors/Typos*
- *Equation Display*

We kindly request that you note your revisions directly in the PDF provided to you using the "comment function" in Adobe Reader. You can find a quick tutorial on how to use this function attached to your PDF. Should you choose to submit your revisions in any other manner, IGI Global reserves the right to send the book to print without your requested changes. Please note that this requested format for revisions allows IGI Global's book production staff to better address your revisions and also allows this book to be sent to print in a timely

IGI Editorial:

Lindsay Johnston, Managing Director
Christina Henning, Production Editor
Austin DeMarco, Managing Editor
Jeff Snyder, Copy Editor
Sean Eckman, Development Editor
Henry Ulrich, Production Assistant

IGI PUBLISHING
WWW.IGI-GLOBAL.COM

International Journal of People-Oriented Programming

July-December 2014, Vol. 3, No. 2

Table of Contents

The *International Journal of People-Oriented Programming* is indexed or listed in the following: Bacon's Media Directory; Cabell's Directories; INSPEC; JournalTOCs; MediaFinder; ProQuest Advanced Technologies & Aerospace Journals; ProQuest Computer Science Journals; ProQuest Illustrata: Technology; ProQuest SciTech Journals; ProQuest Technology Journals; The Index of Information Systems Journals; The Standard Periodical Directory; Ulrich's Periodicals Directory

Understanding Wheel Spinning in the Context of Affective Factors

Ma. Mercedes T. Rodrigo, Ateneo de Manila University, Quezon City, Phillipines

Joseph Barbosa Beck, Worcester Polytechnic Institute, Worcester, MA, USA

ABSTRACT

The notion of wheel spinning, students getting stuck in the mastery learning cycle of an ITS without mastering the skill, is an emerging issue. Although wheel spinning has been analyzed, there has been little work in understanding what factors underlie it, and whether it occurs in cultural contexts outside that of the United States. This work is an extension of an earlier analysis of 116 students in an urban setting in the Philippines. The authors found that Filipino students using the Scatterplot Tutor exhibited wheel spinning behaviors. The authors explored the impact of an intervention, Scooter the Tutor, on wheel spinning behavior and did not find that it had any effect. They also analyzed data from quantitative field observations, and found that wheel spinning is negatively correlated with engaged concentration, positively correlated with confusion, but not correlated with boredom. This result suggests that the problem of wheel spinning is primarily cognitive in nature, and not related to student motivation. However, wheel spinning was positively correlated with gaming the system, and causal analysis suggests that wheel spinning causes gaming.

Keywords: *Affect, Gaming the System, Quantitative Field Observations, Wheel Spinning,*

INTRODUCTION

Mastery learning is a philosophy of teaching and learning that "asserts that under appropriate instructional conditions, virtually all students can learn well, that is, can 'master,' most of what they are taught." The theory further asserts that teachers can deliver instruction so that students learn well, enabling students' short-term success and preparing students for subsequent learning (Block & Burns, 1976). ~~This model is consistent with the intuition that "practice makes perfect."~~

Decades of research have investigated mastery learning in the context of computer-aided instruction (see Corbett & Anderson, 1994; Frick, 1990). During the 1970s, researchers began developing new modes of computer-based education, specifically intelligent tutoring systems (ITSs). ITSs were learning environments that make use of artificial intelligence to engage students in deep reasoning based on an understanding of students' knowledge and behavior (Corbett, Koedinger, & Anderson, 1997). Many of today's ITSs provide assistance to the student in

DOI: 10.4018/IJPOP.2014070104

the form of hints or by breaking a complex problem down into simpler steps for the student to solve individually.

ITSs are effective learning environments for computer-assisted problem solving because they are capable of providing assistance to students who are confused. Prior work estimates that ITSs are at least as good as human tutoring at helping students learn (VanLehn, 2011).

Although it is tempting to assume all students benefit from using an ITS, this assumption does not necessarily hold. Upon reflection, the mastery learning model has weaknesses: If a student requires assistance to solve his first two problems, presenting a third with the hope he will learn the skill is a reasonable strategy. However, if the student has been unable to reach mastery in the first twenty practice opportunities of a skill, it is unlikely that the twenty-first opportunity will enable the student to acquire the skill (Beck & Rodrigo, 2014). In many ITSs, the student does not need to attempt a fixed number of problems, but continues to solve problems until he achieves mastery of the associated skills. In other words, the computer keeps presenting problem after problem for as long the student's skill is less than mastery level.

Beck and Gong (2013) labeled this phenomenon "wheel spinning". It refers to students who fail to master a skill in a computer tutor in a timely manner. The name comes from a car spinning its wheels in the snow: there is the illusion of forward progress as the car is expending effort and its wheels are spinning rapidly, yet it is not going anywhere. What exactly a "timely manner" is varies from tutor to tutor, but Beck and Gong's (2013) prior analysis shows that, after 10 attempts to practice a skill, approximately two-thirds of students would have achieved mastery. Subsequent attempts barely increase this percentage, implying that the last third of students need other interventions.

This paper is an extension of a previously-published short conference paper by Beck and Rodrigo (2014). This paper further explores the phenomenon of students being stuck on a particular skill, investigates what other constructs relate to it, and discusses possible approaches for remediation. It specifically attempts to address the following research questions:

1. How do students spend their time while using an ITS?
2. Do students wheel spin within the ITS?
3. What is the relationship between affect and wheel spinning?
4. What is the relationship between off-task behavior and wheel spinning?

STUDY POPULATION, SYSTEM, AND METHODOLOGY

In this study of wheel spinning, we measured the relationship between wheel spinning and a variety of affective and cognitive variables. We made use of an existing dataset that was previous analyzed for constructs such as student disengagement (see Rodrigo et al., 2010), help-seeking behavior (see Ogan et al., 2014), affect (see Rodrigo et al., 2011), and carelessness (see San Pedro et al., 2011).

Participants

A public high school in in Quezon City, Philippines permitted the research team to conduct the study on their premises, with their students. As of 2008, the school had 5,368 students, predominantly Filipino, and 216 teachers (*Ramon Magsaysay Cubao High School: School Profile Report*, 2009). The school's community was relatively poor. A survey of the students' parents revealed that about one-half of respondents were unemployed and approximately 70%

of households earned PhP10,000 per month or less (approximately US$230.00). The school had 32-bit Windows XP computers intended for student use, but many were in disrepair or kept in storage and not used for instruction (see Rodrigo et al., 2011). This profile implied that students were studying under relatively resource-poor conditions and had limited prior experience with computers within an educational setting.

The principal and math and science coordinators selected 126 public high school students to participate in the study from different first year high school sections. That is, participants included regular students and advanced math and science students. Students ranged in age from 12 to 14. There were 74 girls and 40 boys. Gender data on the other 2 participants was missing from the data set (see Rodrigo et al., 2011).

Materials

The testbed for this study was the Cognitive Tutor unit on scatterplot generation and interpretation (Baker, et al., 2006). Within the Scatterplot Tutor, the learner is given a problem scenario (Figure 1) and data that he needs to plot (Figure 2) in order to arrive at the solution.

The learner first identifies the variables that each axis of the graph would represent. He then had to provide an appropriate scale for each axis (Figure 3).

He proceeds to label the values of each variable along the axis and plot each of the points of the data set (Figure 4).

Finally, he interprets the resultant graphs.

The Scatterplot Tutor provides contextual hints and supplementary content to guide the learner, feedback on correctness, and messages for errors (Figure 5).

The tutor also monitors and displays the learner's progress through skill bars that depict his mastery of skills (Figure 6). Students work on a total of twenty-four skills within the tutor.

Treatment Conditions

The 126 participants were randomly assigned to treatment condition. Sixth-two participants (control) used the standard Cognitive Tutor lesson on scatter plots. Sixty-four of the participants (experimental) were assigned to use a version of the tutor with an embodied conversational agent, "Scooter the Tutor". Scooter was designed for two purposes. First, Scooter was designed to dissuade a form of off-task behavior called gaming the system. Gaming the system is defined as systematic misuse of system features to advance through the learning materials without actually learning the content (Baker et al., 2004). As shown in Figure 7, Scooter became angry when students gamed the system. Although the goal of this visual was to mitigate ~~gaming, some~~ students actually enjoyed seeing Scooter turn red and would deliberately attempt to get him angry (Ryan Baker, *personal communications*). Second, Scooter was designed to help students learn the material that they were avoiding by gaming, while affecting non-gaming students as minimally as possible. Scooter display~~s~~ happiness and ~~gives~~ a positive message when students ~~do~~ not game (regardless of the correctness of their answers), but ~~shows~~ dissatisfaction when students ~~game~~, and ~~provides~~ supplementary exercises to help them learn material bypassed by gaming (see Figure 5).

Observation Protocols

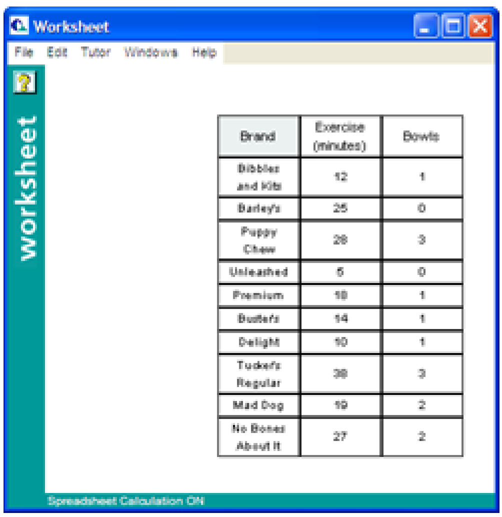
Each student's prevalence of each affective state was assessed using the Baker-Rodrigo Observation Method Protocol (BROMP) (Ocumpaugh, 2012). The protocol is operationalized differently in each field study. To collect the data for this study, prior work (see Rodrigo, et al., 2010; San

Figure 1. Problem scenario example

Samantha is trying to find out what brand of dog food her dog food her dog Champ likes best. Each day, she feeds him a different brand and sees how many bowls he eats. But then her mom says that maybe her dog just eats more on days when he exercises more.

Please draw a scatterplot to show how many bowls the dog eats, given the dog's level of exercise that day

Figure 2. Data that needs to be plotted



Brand	Exercise (minutes)	Bowls
Bibbles and Kits	12	1
Barley's	25	0
Puppy Chew	28	3
Unleashed	6	0
Premium	18	1
Butler's	14	1
Delight	10	1
Tucker's Regular	38	3
Mad Dog	19	2
No Bones About It	27	2

Pedro et al., 2011) reports that each participant was observed 24 times by a pair of trained expert coders, with an interval of 180 seconds between observations lasting 20 seconds each. In order to reduce the degree to which affect is altered by the observation process, observations were conducted using peripheral vision or at an angle where the observer appeared to be looking at another participant, so that the participant currently being observed would not know that he is the one being observed. It was possible for a student to exhibit more than one affective state during a time period. As in past research using this method, however, only the first observed affective state was recorded to avoid bias towards interesting or dramatic events. The coding scheme included seven categories: boredom, confusion, delight, engaged concentration (called flow in Craig, et al., 2004), frustration, surprise and neutral. Observers also recorded the first user behavior from a list of seven behaviors of interest: on-task, on-task giving and receiving answers, other on-task conversation, off-task solitary, other off-task conversation, gaming the

Figure 3. Scaling tool

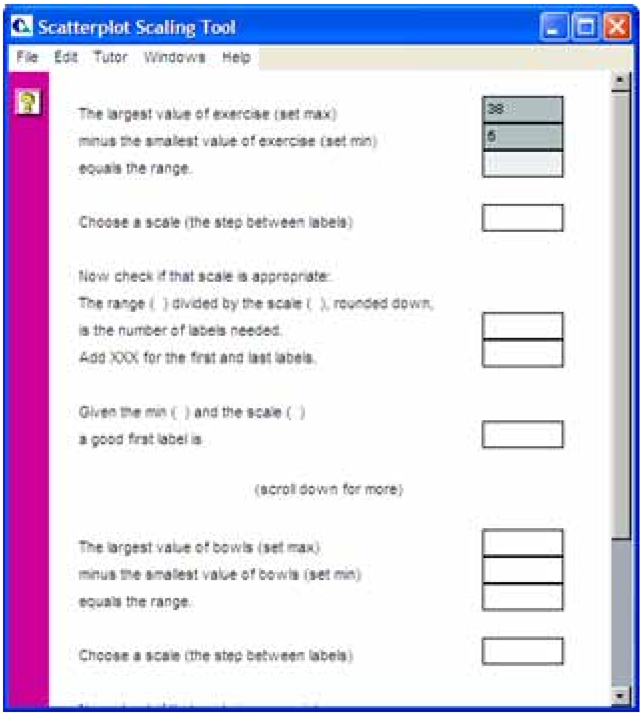


Figure 4. Scatter plot work area

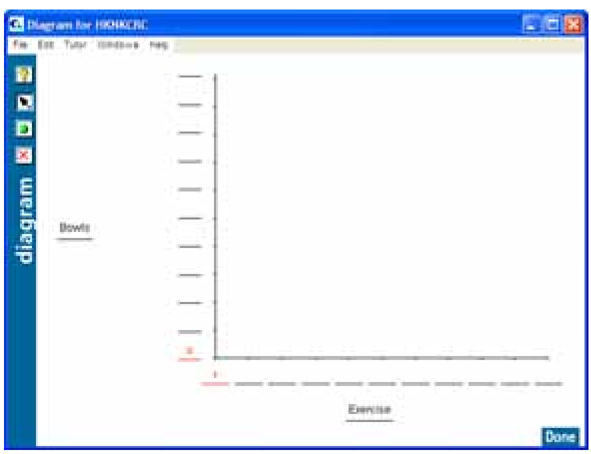


Figure 5. Supplementary content

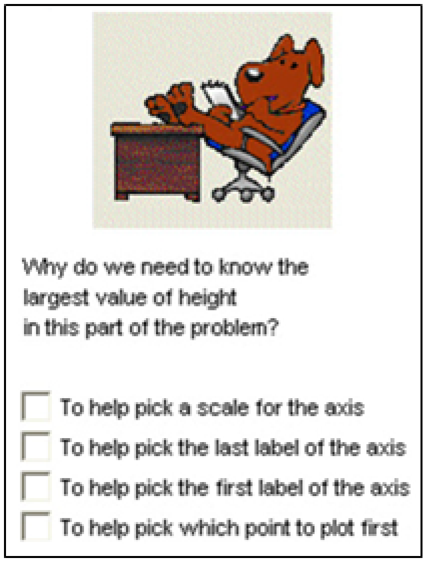


Figure 6. Skill bars

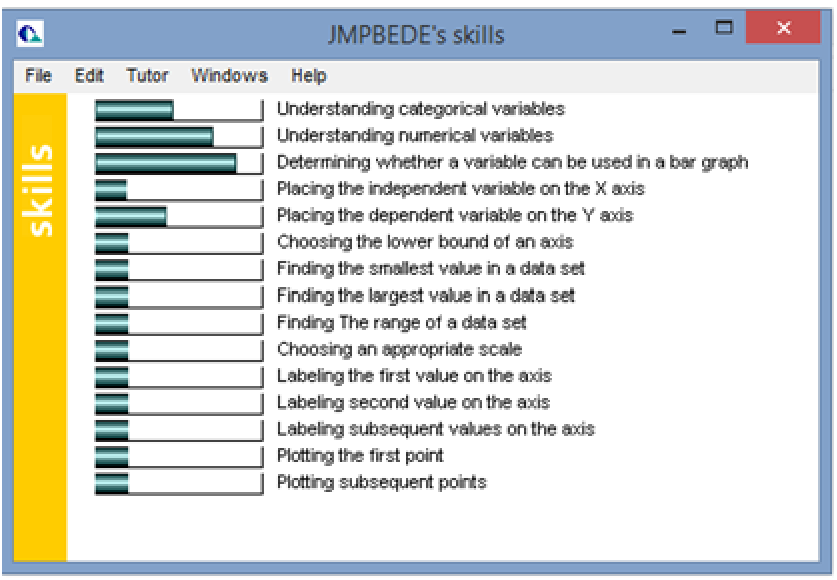


Figure 7. Happy vs. angry Scooter

system, and inactive (Rodrigo, et al., 2011). The observers' inter-rater reliability was found to be acceptable at K=0.54. More details on the quantitative field observation method, including examples from the coding manual can be found in Ocumpaugh (2012).

The participants were organized in groups of 10. Prior to using the software, students viewed a lecture on the topic, delivered via a PowerPoint presentation with a voiceover and some simple animations. This lecture was necessary to level off student knowledge, as students had not explicitly covered these topics in class prior to the study. After the lecture, each group was asked to use the Scatterplot Tutor for 80 minutes. Each student in each group took a nearly isomorphic pre-test and post-test, counterbalanced across conditions.

We started with 126 students in our study population. Unfortunately, data from four students were lost due to a faulty USB drive. In addition, there were problems reconciling affect observations with interaction data for six students. Therefore, we were left with 116 students to analyze. Furthermore, we noticed that some students solved relatively few problems in the Scatterplot Tutor, resulting in two difficulties. First, such students worked on fewer different topics, and thus saw a subset of the material other students saw. Second, these students' affective states were based on few observations, resulting in statistically unstable estimates of the affective measures. Therefore, for analyses involving affect, we screened out students who completed fewer than 60 problems. After this attrition, we were left with 57 students in the control group and 49 students in the experimental group.

RESULTS

Investigating Student Mastery in the Scatterplot Tutor

In order to investigate how students mastered content in the Scatterplot Tutor, we made use of the log files recorded during the study to analyze student performance. The analyses that follow use the entire dataset, i.e. both the control and experimental group data.

How ~~did~~ students spend their time in the Scatterplot Tutor? As we reported earlier (Beck & Rodrigo, 2013), we separated students into three categories of learners on any given problem. The first category was composed those students still working towards mastery. The second category was composed of those who had just mastered the skill on that problem. The third category was composed of those learners who had mastered the skill on a previous problem. For the purposes of this paper, we use a definition of mastery defined as three correct responses in a row. Figure 8 shows how many students were engaged in each of these three activities for the first 20 practice opportunities of each skill. The y-axis of the graph goes up to 2610, since there are 24 skills in the Scatterplot Tutor, and 116 students (some students did not attempt all of the skills). Therefore, on the first practice opportunity, all students were working towards mastering the skill, as none could have mastered it yet (since the definition is three correct responses in a row). On the third practice opportunity, a fair proportion of the students mastered the skill. By the seventh practice opportunity, relatively few students were still working towards mastery, and those students were unlikely to master the skill. The majority of students were working on additional practice of the skills, and possibly overpracticing (Cen, Koedinger, & Junker, 2007). Whether all of this overpractice is wasted or even preventable (~~FAnesali~~, Nixon, & Ritter, 2013) is debatable; however we were surprised at the low number of students, both in absolute terms and as a relative proportion, still working towards mastering the skill by the 9^{th} practice opportunity.

Do students wheel spin within the scatterplot tutor? To investigate whether wheel spinning occurs within the Scatterplot Tutor, we graphed the number of practice opportunities on the x-axis, and on the y-axis plotted the percentage of students who succeeded in mastering the skill. Figure 8 shows performance of students within the Scatterplot Tutor. The interpretation for this graph is that if you wait until the 20^{th} practice opportunity, what percentage of students will have mastered the skill?

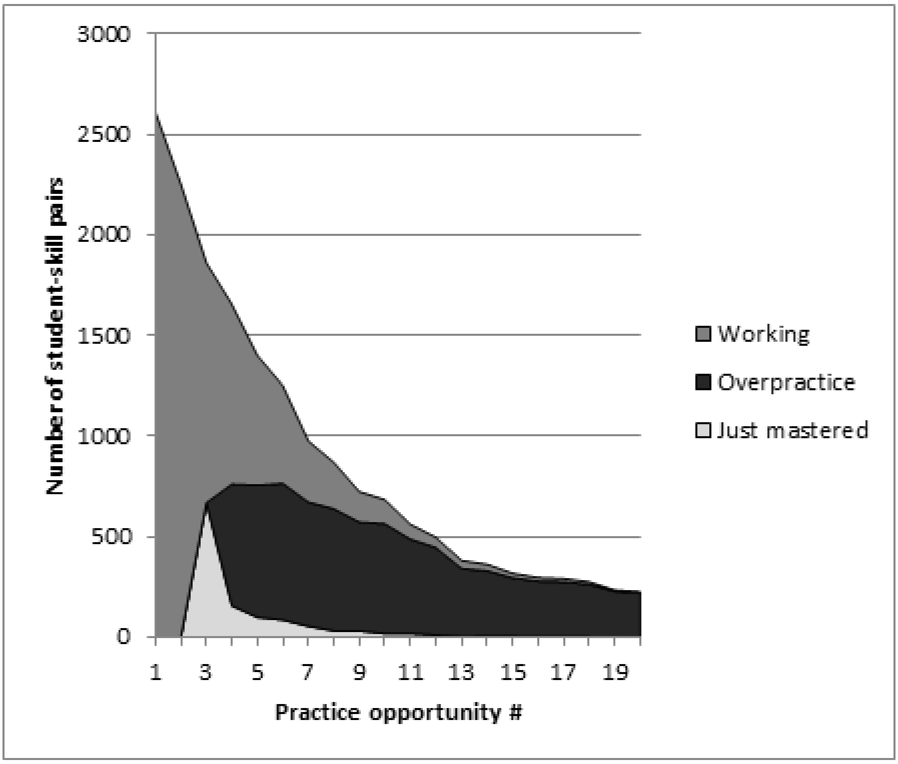
The number of students who have mastered a skill is easy to compute: count the number of learners who have so far succeeded in answering three questions in a row correctly. These are the students who have mastered the skill. To compute a percentage, we can simply compute (#mastering-the-skill / #attempting-the-skill). Unfortunately, there are two defensible choices for the denominator, which we will now explore. The first candidate is the total number of students who have started the skill and completed one practice opportunity. This approach corresponds to the solid, lower, line in Figure 9, and is consistent with prior work on wheel spinning (Gong & Beck, 2013). However, this approach causes an artificially negative view of student performance, as a learner who only solves 1 problem has no chance of mastering the skill (as the definition requires 3 correct responses in a row), but still counts towards students who failed to achieve mastery. Although this student failed to achieve mastery, it is not necessarily fair to count this failure against the tutor.

An alternate candidate for the denominator is to only count students who have attempted at least 3 problems for the skill. This way, those students have the opportunity to master the skill, and as seen in Figure 9, many students master the skill at this point. Using the more restrictive definition of only counting students who have attempted 3 or more problems does not affect the numerator, but shrinks the denominator, resulting in a higher percentage of students mastering the skill. This definition of deciding who has attempted a skill corresponds to the upper, dashed line. One drawback of this definition is that a student may begin a skill, attempt to solve one or two problems, and then decide the material is too hard and give up. These students could be considered as failing to achieve mastery, but are not accounted for by the dashed line.

While neither definition of students attempting a skill is perfect, they give some notion of the scope of the problem. If we only considered students who attempted 3 or more problems, and waited until the 15^{th} practice opportunity, only 63.2% of students mastered the skill. Waiting until 20 practice opportunities resulted in very few additional students mastering the skill (63.4%), as relatively few students attempted to solve that many problems. In short, students who mastered the material in the Scatterplot Tutor tended to do so quickly; after 7 practice opportunities 90% of the students who will eventually master the skill have already done so. So while computer tutors are able to provide infinite practice, few students have the patience to attempt more than 10 problems.

These results, combined with Figure 8, which demonstrated that relatively few students succeed in mastering a skill relative to the number working on it, suggest that Filipino students working in the Scatterplot Tutor were capable of exhibiting wheel spinning behavior. After a student attempted 10 problems on a skill and failed to master the skill, he had little hope of

Figure 8. Number of students engaged in each type of activity as a function of practice opportunity



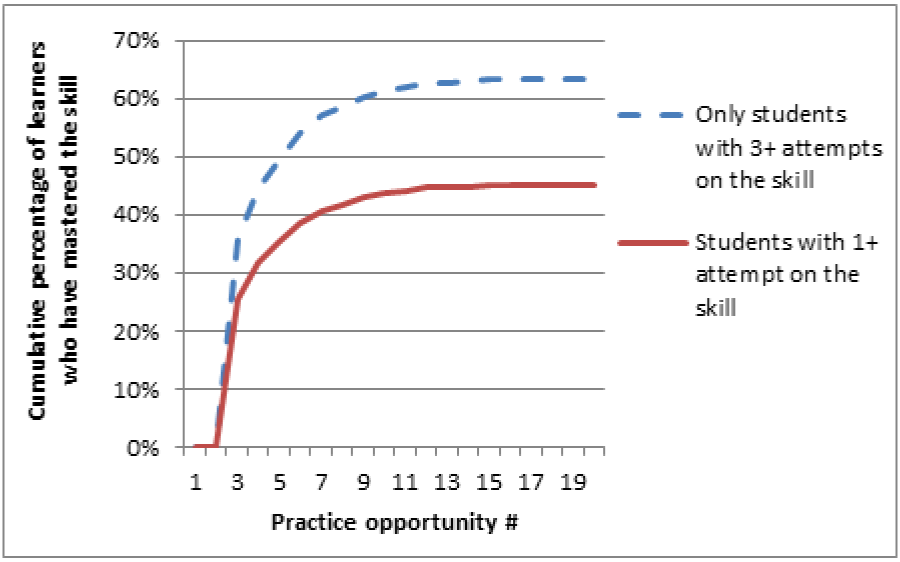
achieving mastery through additional interaction with the ITS. For consistency with prior research, we also adopt a threshold of 10 problems for our cutpoint for wheel spinning. That is, if a student reaches 10 problem attempts on a skill without mastery, we define him as exhibiting wheel spinning behavior on that skill (Beck & Rodrigo, 2013).

UNDERSTANDING THE INTERPLAY OF SCOOTER THE TUTOR, AFFECT, AND WHEEL SPINNING

For interpreting the affect data, in order to obtain scientifically meaningful results, we restricted the data in two ways. First, as mentioned previously, we excluded students who solved fewer than 60 problems. Second, we found that certain affective states were rarely observed by our coders. Table 1 shows how often students were coded as being in each affective state at least once. Thus, all students were at times in a state of confusion and engaged concentration. Conversely, only 5% of students were noted to observe surprise at least once. This lack of observations for certain affective states makes it difficult to draw statistical conclusions about those states from our data (Beck & Rodrigo, 2013).

According to Table 1, only 9% of students were observed to be frustrated while using the Scatterplot Tutor. There are several possible reasons for this low number:

Figure 9. Student mastery vs. practice opportunity



1. The Scatterplot Tutor could ~~be well designed to minimize~~ frustration, and the 9% number ~~is~~ an accurate assessment.
2. Students may ~~be~~ skilled at hiding their emotions, particularly difficulty in learning which may ~~carry~~ ~~negative~~ social stigma.
3. The human coders might have interpreted frustration as another state.
4. Students may have exhibited frustration when they were not under observation.

Regardless of the cause, states that were rarely observed will present problems for statistical analyses. When 95% of the values in a dataset are 0, statistical tests will be hard-pressed to detect effects and violations of normality are a given. Given the lack of statistical power, and extreme non-normality of the data, we did not examine the affect states of ~~None, Frustration, or Surprise. That left us with Confusion, Engaged concentration, Boredom, Neutral, and Delight,~~ as well as our measure of percent time gaming the system and percent of skills on which the student wheel spun.

Both the control and experimental groups worked with the Scatterplot Tutor. In addition, the experimental group received feedback and assistance from Scooter the Tutor. One question is whether Scooter had an impact on the affective states or on the amount of wheel spinning. The impact of Scooter on affective states has been previously studied (Rodrigo, et al., 2012), and this work replicates the finding that the presence or absence of Scooter did not result in a statistically reliable difference between conditions. We also measured Scooter's impact on wheel spinning. In both conditions, the mean of wheel spinning was 0.37 so there appeared to be no impact from Scooter. Given that Scooter also included instruction, it is somewhat surprising that Scooter did not affect the rate of wheel spinning. To further explore wheel spinning, we examined how the other constructs we measured correlated with it.

What is the interrelationship between affect and wheel spinning? As we expected student initial knowledge to directly affect both wheel spinning and affective measures such as confusion and engaged concentration, we computed partial correlations, partialing out the student's

Table 1. Percentage of students exhibiting various affective states

Affective State	Percentage of Students Observed in That State at Least Once
Confusion	100%
Engaged concentration	100%
Boredom	33%
Neutral	29%
Delight	25%
None	14%
Frustration	9%
Surprise	5%

pretest score (Beck & Rodrigo, 2013). Table 2 provides the results of the partial correlations. Wheel spinning was moderately related to engaged concentration and confusion. As expected, more observations of engaged concentration indicated that a student was less likely to wheel spin, while more observations of confusion indicated a student was more likely to wheel spin. There was also a moderately strong relationship with gaming the system, with students who gamed more being more likely to wheel spin. Although the direction of the relationships was unexpected, the magnitude of the correlation was not. Given that prior domain knowledge was controlled for, we were surprised at the strength of the relationships and expected much weaker associations. Perhaps most interestingly, boredom was not strongly related to wheel spinning, with a partial correlation of 0.15 (Beck & Rodrigo, 2013).

In addition to correlating wheel spinning with students' overall incidence of engaged concentration, confusion, gaming, boredom, and delight, we examined the relationship between wheel spinning and persistent confusion and persistent engaged concentration. We operationalized persistent confusion or engaged concentration as three consecutive observations of the state. Since observations were 3 minutes apart, that would be estimates of confusion or concentration over a 9-minute interval. We counted the number of times each student was observed to be confused-confused-confused and engaged concentration-engaged concentration –engaged concentration. We divided these counts by the total number of three-observation sequences to get the incidence of these two sequences per student. We chose these two constructs to use for persistence as they had the most observations. The next most frequently occurring affective state, boredom, had only 10% of students exhibiting even one instance of repeated boredom, an insufficient sample from which to compute a reliable correlation. In contrast, engaged concentration and confusion had instances of persistent behavior in 80% and 88% of students, respectively.

Our reason for checking persistence was that we suspected that students who were persistently confused would be more likely to wheel spin than students who were only intermittently confused, even if overall both students had the same average level of confusion. For example, a student who is confused three times at widely separated points in time is probably exhibiting less wheel spinning than a student who is consistently confused over a short period of time. Contrary to our expectations, the correlations between wheel spinning and persistent confusion and engaged concentration were somewhat weaker than the correlations with the overall levels of confusion and concentration.

Table 2. Partial correlations of immediate affect variables vs. wheel spinning controlling for pretest score

Construct	Partial Correlation	P-Value
Overall (N=106)		
Flow	-0.52	1.03×10^{-8}
Confusion	0.48	2.91×10^{-7}
Gaming the system	0.44	3.24×10^{-6}
Boredom	0.15	0.14
Delight	0.05	0.59
Confusion-Confusion-Confusion	0.36	6.8×10^{-5}
Flow-Flow-Flow	-0.46	2.1×10^{-7}
Control (N=57)		
Flow	-0.46	0.0004
Confusion	0.46	0.0004
Gaming	0.30	0.025
Boredom	0.03	0.85
Delight	0.09	0.51
Confusion-Confusion-Confusion	0.492	0.0001
Flow-Flow-Flow	-0.37	0.005
Experimental (N=49)		
Flow	-0.61	4.8×10^{-8}
Confusion	0.49	0.0004
Gaming	0.58	0.00003
Boredom	0.25	0.08
Delight	0.014	0.92
Confusion-Confusion-Confusion	0.41	0.004
Flow-Flow-Flow	-0.55	0.00005

Table 2 also provides the relationship between our affective measures and wheel spinning broken down by whether students were in the control or experimental version of the tutor. There have been observations of students changing behavior in response to Scooter, so we wondered whether such changes would alter the relationship between observed affect and wheel spinning. In general, the magnitude of the partial correlations was noticeably stronger for the experimental than for the control condition. One possible explanation is that Scooter also provided a tutoring component and attempted to discourage students from gaming. Students who exhibted confusion in spite of the extra tutoring were presumably struggling even more and are thus more likely to wheel spin. The increase in magnitude of partial correlations extends to the results for gaming (0.30 for control vs. 0.58 for experimental). One possible explanation is that students in the experimental group received visual feedback as a result of gaming (see Figure 7, discussed previously), and so consequently, may have done gaming actions with the goal of angering Scooter.

Such actions would result in an increase in wheel spinning and would not be accounted for by the partial correlate of pretest score.

Combined, these results suggest that students are wheel spinning not because of affective factors where they are not motivated to do the work, but rather, students are genuinely stuck on the material and need additional instructional support. To test this intuition formally, we modeled the problem in Tetrad, a freely available tool for causal discovery in datasets[1] (Beck & Rodrigo, 2013). Tetrad takes advantages of asymmetries in partial correlations between variables to, sometimes, discover which variables are causal (Pearl, 2000) rather than merely associated. Briefly, Tetrad tests whether a variable, C, is able to partial out the relationship between A and B. If so, A and B are (with extremely high probability) not causally related (Pearl, 2000).

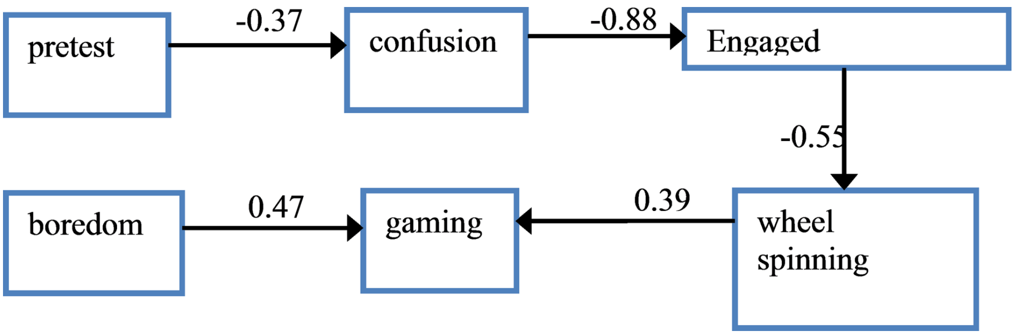
Aside from causal discovery, the graphs Tetrad produces provide a succinct representation of partial correlations and direct relationships between variables. Due to problems of multiple hypothesis testing, we restricted our analysis to only consider confusion, engaged concentration, and boredom as these variables were the most related to wheel spinning, and they were also the states that had the most observations by the human coders. In addition, we included domain pretest score, wheel spinning, and gaming the system in the model.

We used the Tetrad's PC search algorithm to discover the structure, and its estimator functionality to estimate the model coefficients. We first Z-transformed the data to normalize it, and consequently make the coefficient magnitudes comparable. In addition, we provided Tetrad with background knowledge that the pretest score was causally upstream from all of the other variables. That is, pretest score *could* be a cause of the other variables, but could not be caused by any of the variables collected while students were interacting with the tutor (e.g., wheel spinning or engaged concentration). Figure 10 provides the result of our analysis within Tetrad.

The interpretation of Figure 10 is that an arrow from one node to another means there was a *direct causal* relation between the two (Beck & Rodrigo, 2013). For example, from Table 2 we know that confusion and wheel spinning are correlated with each other. However, Tetrad's search determined that patialling out engaged concentration removed the relation between confusion and wheel spinning, and so there was no arrow connecting those two states. Therefore, an intervention that reduced confusion would result in an increase in engaged concentration, and, consequently, the increase in engaged concentration would reduce wheel spinning.

In addition to the edges, there are also associated weights providing the strength of the relationship, and its direction. There are several interesting implications from Figure ~~9.~~ First, Tetrad's search agreed that wheel spinning was related to cognitive factors such as prior knowledge (pretest score), confusion, and engaged concentration, but not to boredom (Beck & Rodrigo, 2013). That is, an intervention targeting boredom is unlikely to affect wheel spinning, while one targeting one of the three causally upstream variables is likely to succeed. Second, the search suggests that gaming was causally downstream of wheel spinning, and was a function of both affective (boredom) and cognitive (confusion / lack of knowledge) factors. This analysis of course is limited by the statistical power of the dataset, and by the variables entered into the analysis (Beck & Rodrigo, 2013). For example, there could be other variables, which we neglected to measure, that could alter the output of Tetrad. This behavior is not a weakness of the tool, but is a fact of the scientific method: we color our interpretations by the data collected. That said, our measures of prior knowledge, affect, within-tutor performance, and gaming, are in broad agreement of what other researchers in the field are gathering.

Figure 10. Causal model of wheel spinning, gaming, and affective states (Beck & Rodrigo, 2013)



CONTRIBUTIONS, FUTURE WORK, AND CONCLUSIONS

This work advances the state of knowledge of the field in several ways. First, it places the phenomenon of wheel spinning in a broader research context. Prior work was restricted to investigating students in the United States (Beck & Gong, 2013), while this work establishes that wheel spinning occurs in at least one non-Western population (Beck & Rodrigo, 2013). Most prior work in ITS has focused on the happy part of the mastery learning cycle: the student masters a skill and stops practice. Recently there has been interest in the phenomenon of overpractice of a skill. This work examines overpractice and wheel spinning, and finds that there are potentially many more students engaged in overpractice than are making progress towards mastery.

This work also examined factors that could influence the rate of wheel spinning. This work replicates and extends prior research linking gaming and wheel spinning (Beck & Gong, 2013). The prior research used a custom-built gaming detector that had not been well validated (Gong, et al., 2010). This work uses a well-validated detector of gaming (Baker, et al, 2004) with broadly similar results in that wheel spinning and gaming appear to be linked (Beck & Rodrigo, 2013). In addition, the direction of causality between gaming the system and wheel spinning was unclear. This work presents evidence that wheel spinning is caused by a deficit in student knowledge and not by student boredom. Gaming the system is caused by wheel spinning or by boredom. That is, students engage in gaming behaviors when they are unlikely to be able to complete the lesson, or are bored and do not care.

Third, this work investigated whether a tutorial intervention, Scooter the Tutor, could influence the amount of wheel spinning. Scooter addresses behavioral issues, as he is triggered by gaming behavior, as well as cognitive deficits through his instructional lessons. Although wheel spinning is related to cognitive deficits, Scooter was not found to be an effective intervention in this study for wheel spinning (Beck & Rodrigo, 2013). To be fair, Scooter was never designed for wheel spinning and does, in fact, succeed at mitigating gaming, the behavior it was designed to address (Baker et al., 2006).

This work also investigated the relationship between affect and wheel spinning. Wheel spinning was found to be related to states that have strongly cognitive natures, i.e. it was negatively correlated with engaged concentration and positively correlated with confusion. On the other hand, boredom and delight did not have any strong correlations with wheel spinning. These findings suggest that students wheel spin because they are stuck and do not know what to do—not because they are bored.

There are several interesting next steps to take from this work. One avenue is to find an intervention that is capable of affecting the rate of wheel spinning. It would also be interesting to perform a fuller analysis of how wheel spinning relates to affective states. For this study, we were limited by the low rates of frustration and surprise in the set of analyses we were able to conduct. In particular, we suspect frustration and wheel spinning are related (Beck & Rodrigo, 2013).

The other major limitation of this work was the lack of temporal linkage between wheel spinning and the affective variables. We had periodic observations of affect, which were sufficient to detect overall the overall level of each state for each student. However, these observations were spaced over time, and not possible to relate to how the student was feeling while engaged in wheel spinning. A promising research direction would be to have finer-grained and time-stamped records of student affect to enable linking them to wheel spinning behaviors. For example, learning that frustration tended to occur before wheel spinning would be informative in knowing when to launch an intervention. In addition, such work could shed light on how serious various affective states are. Is student frustration when beginning a problem set a serious issue and indicates the student will become stuck and wheel spin, or is it relatively common and typically resolves itself? At present, we lack sufficient data to draw such inferences. However, answering such questions is a key future direction.

In summary, this paper investigates wheel spinning and affective factors. We have found that wheel spinning exists in non-Western populations, and is related to knowledge deficits rather than student boredom. As a consequence, wheel spinning is best addressed via cognitive, rather than affective, interventions (Beck & Rodrigo, 2013).

ACKNOWLEGMENT

We thank Ryan Shaun Baker, Ma. Ofelia San Pedro, Jenilyn Agapito, Julieta Nabos, Salvador Reyes, Jr., the Ramon Magsaysay Cubao High School, Mrs. Carmela Oracion and the Ateneo Center for Educational Development, the Ateneo Laboratory for the Learning Sciences, the Ateneo de Manila University's Department of Information Systems and Computer Science

REFERENCES

Baker, R. S., Corbett, A. T., Koedinger, K. R., Evenson, S., Roll, I., Wagner, A. Z., & Beck, J. E. et al. (2006). Adapting to when students game an intelligent tutoring system. In *Intelligent tutoring systems* (pp. 392–401). Springer Berlin Heidelberg. doi:10.1007/11774303_39

Baker, R. S., Corbett, A. T., Koedinger, K. R., & Wagner, A. Z. (2004, April). Off-task behavior in the cognitive tutor classroom: when students game the system. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, (pp. 383-390) doi:10.1145/985692.985741

Beck, J., & Rodrigo, M. M. T. (2014). Understanding wheel spinning in the context of affective factors. In *Intelligent Tutoring Systems* (pp. 162–167). Springer International Publishing. doi:10.1007/978-3-319-07221-0_20

Beck, J. E., & Gong, Y. (2013). Wheel-spinning: Students who fail to master a skill. In Artificial Intelligence in Education. Springer Berlin Heidelberg, (pp. 431-440).

Block, J. H., & Burns, R. B. (1976). Mastery learning. *Review of Research in Education*, ▪▪▪, 3–49.

Cen, H., Koedinger, K. R., & Junker, B. (2007). Is Over Practice Necessary?-Improving Learning Efficiency with the Cognitive Tutor through Educational Data Mining, In Frontiers in Artificial Intelligence and Applications, 158, 511.

Corbett, A. T., & Anderson, J. R. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction*, *4*(4), 253–278. doi:10.1007/BF01099821

Corbett, A. T., Koedinger, K. R., & Anderson, J. R. (1997). Intelligent tutoring systems. Handbook of human computer interaction, 849-874.

Craig, S., Graesser, A., 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. doi:10.1080/1358165042000283101

Fancsali, S. E., Nixon, T., & Ritter, S. (2013). Optimal and worst-case performance of mastery learning assessment with Bayesian Knowledge Tracing. In *Proceedings of the 6th International Conference on Educational Data Mining*.

Frick, T. W. (1990). A comparison of three decision models for adapting the length of computer-based mastery tests. *Journal of Educational Computing Research*, *6*(4), 479–513. doi:10.2190/REB5-92JC-YPQM-AMTE

Gong, Y., Beck, J., Heffernan, N. T., & Forbes-Summers, E. (2010). The impact of gaming (?) on learning at the fine-grained level. In *Proceedings of the 10th International Conference on Intelligent Tutoring Systems (ITS2010)*, (Vol. 1, pp. 194-203) doi:10.1007/978-3-642-13388-6_24

Ocumpaugh, J. (2012). Baker-Rodrigo observation method protocol (BROMP) 1.0. Training manual version 1.0. *Baker-Rodrigo Observation Method Protocol (BROMP) 1.0. Training Manual version 1.0, Tech. Rep*.

Ogan, A., Walker, E., Baker, R., Rodrigo, M. M. T., Soriano, J. C., & Castro, M. J. (2014). Towards understanding how to assess help-seeking behavior across cultures. *International Journal of Artificial Intelligence in Education*. doi: doi:10.1007/s40593-014-0034-8

Ramon Magsaysay Cubao High School: School Profile Report. 2009, Ateneo de Manila University, Loyola Heights, Quezon City, Philippines.

Rodrigo, M. M. T. d Baker, R. S., Agapito, J., Nabos, J., Repalam, M. C., & Reyes Jr, S. S. (2010, January). Comparing disengaged behavior within a cognitive tutor in the USA and philippines. In Intelligent Tutoring Systems, Springer Berlin Heidelberg, (pp. 263-265).

Rodrigo, M. M. T., Baker, R. S. J. d, Agapito, J., Nabos, J., Repalam, M. C., Reyes, S. Jr., & San Pedro, M. O. C. Z.. (2011). The effects of an embodied conversational agent on student affective dynamics while using an intelligent tutoring system. *IEEE Transactions on Affective Computing, 2*(4), 18-37

San Pedro, M. O. C. Z., Rodrigo, M. M. T., & Baker, R. S. J. D. (2011). The relationship between carelessness and affect in a cognitive tutor. In S. D'Mello & A. Graesser (Eds.), *ACII 2011, Part I, LNCS 6974* (pp. 306–315). Berlin, Heidelberg: Springer-Verlag. doi:10.1007/978-3-642-24600-5_34

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, *46*(4), 197–221. doi:10.1080/00461520.2011.611369

ENDNOTES

[1] http://www.phil.cmu.edu/projects/tetrad/