 ISSN: 2167-1435
v31 n3086

# Examining Student Learning in Spreadsheet Assignments: The value of activity-trace logs

Gove Allen
gove@byu.edu
Information System Department
Brigham Young University
Provo, UT 84602, USA

Nicholas Ball
nicholas.ball@uvu.edu
Information Technology Department
Utah Valley University
Orem, UT 84058, USA

John Chapman
chapmanjs@byu.edu

Randy Davies
randy.davies@byu.edu

Instructional Psychology and Technology Department,
Brigham Young University
Provo, UT 84602, USA

## Abstract

In most problem-solving assignments, professors evaluate student solutions without the ability to observe the process students used to arrive at their solutions. This paper presents an approach for allowing professors to have detailed, activity-trace process data about how students arrived at solutions, giving insights into reasoning and student misunderstandings that happen (and are sometimes corrected) prior to submission. By rendering the assignment in Excel and using a template configured to log cell changes, the files submitted by students contain transactional level data for each attempt made. These activity-trace logs also provide a powerful mechanism to tell when students are copying work from other students. An example of how this can help instructors understand the scope of student misconception is presented.

**Keywords:** Excel, Educational Data Mining, Learning Analytics, data logging, analyzing student learning.

©2014 EDSIG (Education Special Interest Group of the AITP), www.aitp-edsig.org
and FITE (Foundation for Information Technology Education), edfoundation.org/ 

## 1. INTRODUCTION

In problem-solving homework exercises, students are often given a task and required to produce a solution. Typically, the student's solution is used to evaluate mastery of the subject matter and assign a grade. This approach limits what a professor can know about student understanding or lack thereof. Even in assignments where students must show their work as is commonly done in mathematics education, the professor only has access to the logical steps that the student presents as leading from the initial state (the problem description) to the goal state (the presented solution). They do not have access to the efforts a student makes in coming to that final solution.

## 2. OBSERVING PROCESS

A much richer understanding of student comprehension can be gained by examining in detail the problem-solving process that a student engages in as he or she develops a solution (Behrens, Mislevy, DiCerbo, & Levy, 2011). In fact, this is exactly what happens when a struggling students seeks consultation with a professor. The professor might have a student work through a problem analyzing the process undertaken by the student. As the professor does this he or she gains insight into any misunderstanding (i.e., knowledge gaps and misconceptions) the student may have about the problem. This process informs any remedial action that is required and hopefully helps the student develop a more complete mental model of how to solve a particular class of problems.

While this one-on-one process can be extremely beneficial for the student, it is impractical for the professor to conduct on a broad scale (Martinez-Maldonado, Clayphan, Yacef, & Kay, 2014; Siler & VanLehn, 2014; VanLehn, 2011). The traditional solution has been to guide students to teaching assistants or tutors who can spend individualized time to help students. This approach has at least five drawbacks that prevent it from being a complete solution. First, because the student and the tutor must be collocated, it is inconvenient for the student to meet to get the needed help. While communication technologies (such as Skype and screen sharing software) have been used to allow virtual meetings in similar contexts, the need for the tutor to observe the problem-solving process in detail, reduces the efficacy of the tutoring experience in electronically-mediated interactions. Even in cases where electronically mediated tutoring can be effective, it still requires temporal synchronicity that may prove a barrier as students' schedules often relegate study time to hours when tutors are unlikely to be available.

The second obstacle is the time required for a tutor to be able to observe how and when a student is struggling (Siler & VanLehn, 2014). The process of solving a particular problem may be complicated and because the tutor often has no prior knowledge of the student's misconceptions, he or she must observe the student's progress from the beginning as part of the diagnostic process. It may not be possible for the tutor to "fast forward" the student to the next potentially problematic step in a particular problem and must instead wait as the student plods through the solution. This obstacle of required time manifests itself as two very direct costs. As the limited resource (the tutor) becomes a bottleneck and a queue develops, students incur a cost in the form of time spent waiting to get help. A second cost is often realized in trying to reduce the first: more tutors are hired at a financial cost to either the educational institution or to the student. Because matching the supply of tutor time to student demand is not a simple problem itself, all too often, both costs are substantial.

The third problem with employing tutors to facilitate student learning is one of decentralization and feedback (Aleven, McLaren, Roll, & Koedinger, 2004). When multiple tutors are addressing deficiencies in student understanding, it is more difficult for any individual to identify common patterns among those deficiencies. Accordingly, the likelihood that a misunderstanding will be identified as a common one is substantially reduced. Once a professor is aware of a common misunderstanding, he or she is in a position to address the misunderstanding in lecture, benefitting future students if not current ones. However, because decentralized tutoring reduces both the likelihood that a common misunderstanding will be identified and the likelihood that (once identified) it will be communicated to the professor, the chances that an appropriate curricular adjustment will be made is greatly reduced.

ISSN: 2167-1435
v31 n3086

The fourth problem is that for a student to be assisted in overcoming his or her lack of knowledge, two conditions must exist (Sewell, 2002). First, he or she must be aware of the misunderstanding or at least recognize that there is some problem. Second, he or she must be willing to bear the cost to seek help. Often, when students labors under a misunderstanding, they are unaware that there is a problem. In fact, the term "misunderstanding" connotes that something is understood incorrectly; it stands apart from "confusion" which indicates that something is not understood. Accordingly, students are often unaware of misunderstandings until they have held them for some time and receive feedback on assignments where their performance is less than expected. Even then, they may not be willing to bear the cost to seek help from a tutor. In fact, they will likely exhaust other lower-cost techniques before turning to a tutor such as reading the text, searching the internet, or talking to a fellow student. While any of these may be appropriate and have the effect of helping the student to overcome the misunderstanding, none provide any information to the professor that the misunderstanding existed or how long it persisted.

A final problem in relying on tutors to help student identify and resolve misunderstandings is that for certain classes, there may not be a ready supply of individuals capable of filling the tutor role (VanLehn, 2011). This is the case when the particular course is taken largely by students just prior to graduation.

## 3. THE PEDAGOGICAL ROLE OF ANALYTICS

As the Internet has affected so many aspects of business and electronic transactions have become commonplace, the systematic analysis of data surrounding these transactions has increased dramatically (Behrens, Mislevy, DiCerbo, & Levy, 2011; Johnson, Adams, & Cummins, 2012). These analytics inform virtually every aspect of business from product development, to supply chain management, to advertising and sales. In education, as in business, analytics have developed largely around areas where data can be gathered easily. For learning analytics, this has meant that the focus has been on things like the amount of time a student spends engaged with the electronic learning resources provided to them (e.g. online textbooks and videos). Other easily obtained data included how well students perform on assigned tasks and assessments (e.g. which questions they get right and wrong). Analyses of these data have led to some insights in student learning; however, they generally fall short when used to determine the extent to which students have fully mastered the learning expected of them or used as formative input into remedial actions needed to correct misunderstandings (Behrens, Mislevy, DiCerbo, & Levy, 2011; Siemens, 2012; VanLehn, 2011).

The quality and type of data educators gather often determines the effectiveness of the decisions they make. Chung (2014) describes three levels of data that might be collected for the purpose of informing decisions about a student's performance. System level data is the highest most general form of data that can be collected. System level data might include a profile of students' academic history (e.g., the grades they obtained in the courses they take). Individual level data includes information about a student as they progress through a class. These data might include the results from various assessments or items on a task. The third and deepest level of data we might collect is called transactional data or activity-trace data. Data at the transaction level includes moment to moment actions an individual student takes while completing a task. While many teachers attempt to utilize data to inform instructional decisions, few have adequate access to information at the transactional data level (Siemens, 2012).

## 4. A BETTER APPROACH FOR IDENTIFYING STUDENT MISUNDERSTANDING

Learning analytics is greatly improved if data logging can record student problem-solving progress at the transactional or activity-trace level. It is also best when the collection of data requires little or no time from the professor and no additional time for the student beyond the normal time required to complete homework exercises (Baepler & Murdoch, 2010). It is also best when the review and analysis of the data can centralized and when it can be implemented without the need for students or professors to install specialized software. If these data were available, common errors could be identified through visual inspections as well as through computer analysis.

©2014 EDSIG (Education Special Interest Group of the AITP), www.aitp-edsig.org
and FITE (Foundation for Information Technology Education), edfoundation.org/

 ISSN: 2167-1435
v31 n3086

Using Microsoft Excel as a platform, we have developed such a tool. For any assignment that can be built in such a way that students solve the problem by adding data and formulas to an Excel workbook, our system creates a detailed log of each step the students takes to build a solution not just the final solution graded by the program. The log is recorded on a hidden worksheet within the workbook so that when the student's solution is submitted the log is submitted as well. A student is not even aware that this log exists and it take no addition effort on the students' part. We call the system that builds and maintains these hidden logs the "hidden event log for individual observation system" or HELIOS.

These logs can be aggregated into a single sheet for analysis of students' problem solving success as well as missteps. The consolidated log can also be exported for analysis in other environments including statistical programs or relational database management systems. The tool that aggregates and manages these student logs is called the "activity record evaluation system" or ARES.

We make these two tools, Helios and Ares freely available to professors at accredited institutions of higher education for non-profit, educational use.

## 5. CONFIGURING A HELIOS ASSIGNMENT

The Helios template file comes with a sheet named "Assignment." Figure 1 shows an example assignment configured for use in Helios. Any data located in column B of this sheet will be considered submission data labels and will be collected on a summary page when data from the workbook are aggregated by Ares. The corresponding values to the right of these labels (column C) are recorded for each student workbook when the Ares extracts assignment data and logs from student submissions.

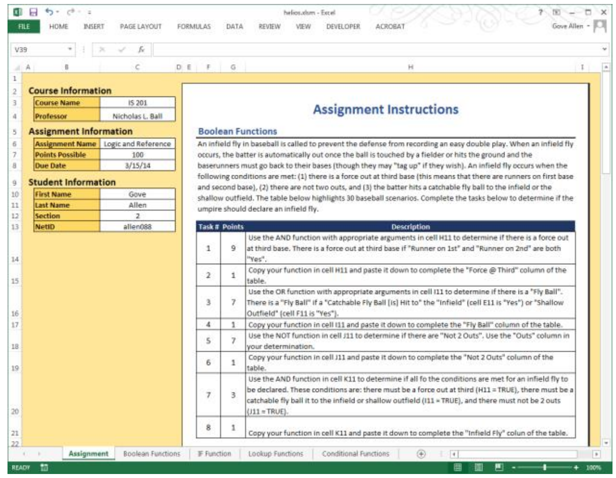
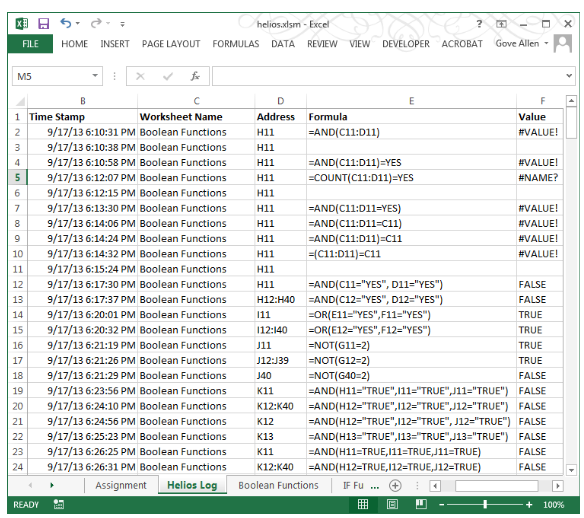



Figure 1: Assignment configured in Helios

## 6. EXAMPLE DATA

We have used the system extensively for teaching Microsoft Excel and have gained several insights into the problem-solving processes of our students. Figure 2 shows log data as recorded by Helios in the student workbook. Although the log is hidden from student view, a new entry is added each time a change is made to a cell. In addition to each change, the time of the change (to the second), the name of the worksheet being modified, the cell label or range of cells being changed, the new content of the cell and the value displayed in the cell at the time of the change are recorded.



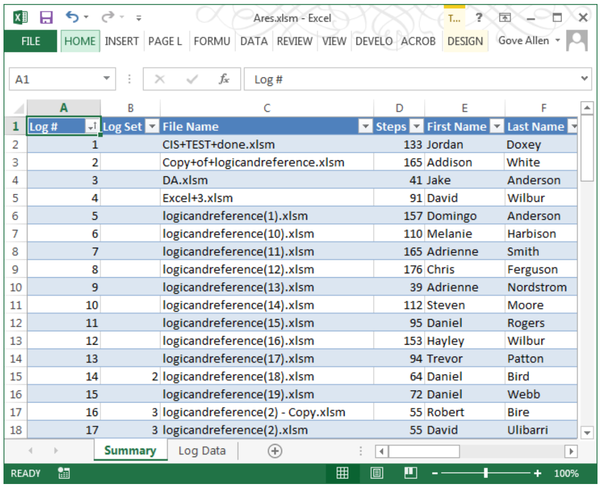
| | Time Stamp | Worksheet Name | Address | Formula | Value |
|---|---|---|---|---|---|
| 2 | 9/17/13 6:10:31 PM | Boolean Functions | H11 | =AND(C11:D11) | #VALUE! |
| 3 | 9/17/13 6:10:38 PM | Boolean Functions | H11 | | |
| 4 | 9/17/13 6:10:58 PM | Boolean Functions | H11 | =AND(C11:D11)=YES | #VALUE! |
| 5 | 9/17/13 6:12:07 PM | Boolean Functions | H11 | =COUNT(C11:D11)=YES | #NAME? |
| 6 | 9/17/13 6:12:15 PM | Boolean Functions | H11 | | |
| 7 | 9/17/13 6:13:30 PM | Boolean Functions | H11 | =AND(C11:D11=YES) | #VALUE! |
| 8 | 9/17/13 6:14:06 PM | Boolean Functions | H11 | =AND(C11:D11=C11) | #VALUE! |
| 9 | 9/17/13 6:14:24 PM | Boolean Functions | H11 | =AND(C11:D11)=C11 | #VALUE! |
| 10 | 9/17/13 6:14:32 PM | Boolean Functions | H11 | =(C11:D11)=C11 | #VALUE! |
| 11 | 9/17/13 6:15:24 PM | Boolean Functions | H11 | | |
| 12 | 9/17/13 6:17:30 PM | Boolean Functions | H11 | =AND(C11="YES", D11="YES") | FALSE |
| 13 | 9/17/13 6:17:37 PM | Boolean Functions | H12:H40 | =AND(C12="YES", D12="YES") | FALSE |
| 14 | 9/17/13 6:20:01 PM | Boolean Functions | I11 | =OR(E11="YES",F11="YES") | TRUE |
| 15 | 9/17/13 6:20:32 PM | Boolean Functions | I12:I40 | =OR(E12="YES",F12="YES") | TRUE |
| 16 | 9/17/13 6:21:19 PM | Boolean Functions | J11 | =NOT(G11=2) | TRUE |
| 17 | 9/17/13 6:21:26 PM | Boolean Functions | J12:J39 | =NOT(G12=2) | TRUE |
| 18 | 9/17/13 6:21:29 PM | Boolean Functions | J40 | =NOT(G40=2) | FALSE |
| 19 | 9/17/13 6:23:56 PM | Boolean Functions | K11 | =AND(H11="TRUE",I11="TRUE",J11="TRUE") | FALSE |
| 20 | 9/17/13 6:24:10 PM | Boolean Functions | K12:K40 | =AND(H12="TRUE",I12="TRUE",J12="TRUE") | FALSE |
| 21 | 9/17/13 6:24:56 PM | Boolean Functions | K12 | =AND(H12="TRUE",I12="TRUE", J12="TRUE") | FALSE |
| 22 | 9/17/13 6:25:23 PM | Boolean Functions | K13 | =AND(H13="TRUE",I13="TRUE",J13="TRUE") | FALSE |
| 23 | 9/17/13 6:26:25 PM | Boolean Functions | K11 | =AND(H11=TRUE,I11=TRUE,J11=TRUE) | FALSE |
| 24 | 9/17/13 6:26:31 PM | Boolean Functions | K12:K40 | =AND(H12=TRUE,I12=TRUE,J12=TRUE) | FALSE |

Figure 2: Log Data as recorded

©2014 EDSIG (Education Special Interest Group of the AITP), www.aitp-edsig.org
and FITE (Foundation for Information Technology Education), edfoundation.org/ 

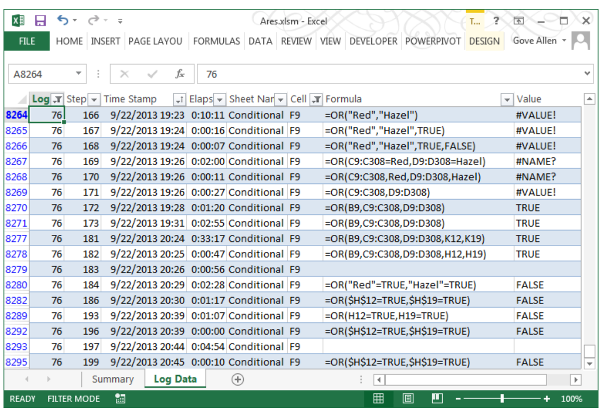
 ISSN: 2167-1435
v31 n3086

Ares is implemented as an Excel workbook that has the capacity to import log files and assignment data from multiple Helios files. Once imported, Ares creates a summary sheet (see Figure 3) as well a combined log with additional fields (see Figure 4).



| | Log # | Log Set | File Name | Steps | First Name | Last Name |
|---|---|---|---|---|---|---|
| 2 | 1 | | CIS+TEST+done.xlsm | 133 | Jordan | Doxey |
| 3 | 2 | | Copy+of+logicandreference.xlsm | 165 | Addison | White |
| 4 | 3 | | DA.xlsm | 41 | Jake | Anderson |
| 5 | 4 | | Excel+3.xlsm | 91 | David | Wilbur |
| 6 | 5 | | logicandreference(1).xlsm | 157 | Domingo | Anderson |
| 7 | 6 | | logicandreference(10).xlsm | 110 | Melanie | Harbison |
| 8 | 7 | | logicandreference(11).xlsm | 165 | Adrienne | Smith |
| 9 | 8 | | logicandreference(12).xlsm | 176 | Chris | Ferguson |
| 10 | 9 | | logicandreference(13).xlsm | 39 | Adrienne | Nordstrom |
| 11 | 10 | | logicandreference(14).xlsm | 112 | Steven | Moore |
| 12 | 11 | | logicandreference(15).xlsm | 95 | Daniel | Rogers |
| 13 | 12 | | logicandreference(16).xlsm | 153 | Hayley | Wilbur |
| 14 | 13 | | logicandreference(17).xlsm | 94 | Trevor | Patton |
| 15 | 14 | 2 | logicandreference(18).xlsm | 64 | Daniel | Bird |
| 16 | 15 | | logicandreference(19).xlsm | 72 | Daniel | Webb |
| 17 | 16 | 3 | logicandreference(2) - Copy.xlsm | 55 | Robert | Bire |
| 18 | 17 | 3 | logicandreference(2).xlsm | 55 | David | Ulibarri |

Figure 3: Ares Summary Sheet

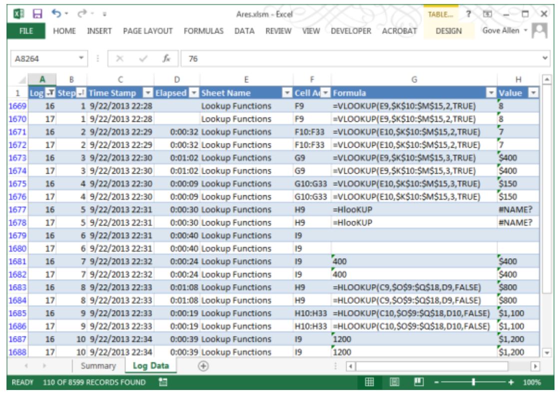


| | Log | Step | Time Stamp | Elaps | Sheet Nar | Cell | Formula | Value |
|---|---|---|---|---|---|---|---|---|
| 8264 | 76 | 166 | 9/22/2013 19:23 | 0:10:11 | Conditional | F9 | =OR("Red","Hazel") | #VALUE! |
| 8265 | 76 | 167 | 9/22/2013 19:24 | 0:00:16 | Conditional | F9 | =OR("Red","Hazel",TRUE) | #VALUE! |
| 8266 | 76 | 168 | 9/22/2013 19:24 | 0:00:07 | Conditional | F9 | =OR("Red","Hazel",TRUE,FALSE) | #VALUE! |
| 8267 | 76 | 169 | 9/22/2013 19:26 | 0:02:00 | Conditional | F9 | =OR(C9:C308=Red,D9:D308=Hazel) | #NAME? |
| 8268 | 76 | 170 | 9/22/2013 19:26 | 0:00:11 | Conditional | F9 | =OR(C9:C308,Red,D9:D308,Hazel) | #NAME? |
| 8269 | 76 | 171 | 9/22/2013 19:26 | 0:00:27 | Conditional | F9 | =OR(C9:C308,D9:D308) | #VALUE! |
| 8270 | 76 | 172 | 9/22/2013 19:28 | 0:01:20 | Conditional | F9 | =OR(B9,C9:C308,D9:D308) | TRUE |
| 8271 | 76 | 173 | 9/22/2013 19:31 | 0:02:55 | Conditional | F9 | =OR(B9,C9:C308,D9:D308) | TRUE |
| 8277 | 76 | 181 | 9/22/2013 20:24 | 0:33:17 | Conditional | F9 | =OR(B9,C9:C308,D9:D308,K12,K19) | TRUE |
| 8278 | 76 | 182 | 9/22/2013 20:25 | 0:00:47 | Conditional | F9 | =OR(B9,C9:C308,D9:D308,H12,H19) | TRUE |
| 8279 | 76 | 183 | 9/22/2013 20:26 | 0:00:56 | Conditional | F9 | | |
| 8280 | 76 | 184 | 9/22/2013 20:29 | 0:02:28 | Conditional | F9 | =OR("Red"=TRUE,"Hazel"=TRUE) | FALSE |
| 8282 | 76 | 186 | 9/22/2013 20:30 | 0:01:17 | Conditional | F9 | =OR($H$12=TRUE,$H$19=TRUE) | FALSE |
| 8289 | 76 | 193 | 9/22/2013 20:39 | 0:01:07 | Conditional | F9 | =OR(H12=TRUE,H19=TRUE) | FALSE |
| 8292 | 76 | 196 | 9/22/2013 20:39 | 0:00:00 | Conditional | F9 | =OR($H$12=TRUE,$H$19=TRUE) | FALSE |
| 8293 | 76 | 197 | 9/22/2013 20:44 | 0:04:54 | Conditional | F9 | | |
| 8295 | 76 | 199 | 9/22/2013 20:45 | 0:00:10 | Conditional | F9 | =OR($H$12=TRUE,$H$19=TRUE) | FALSE |

Figure 4: Partial combined log in Ares

The log shown in Figure 4 has been filtered to show steps 166 through 199 of log number 76 where the student made changes to cell F9. This portion of the log shows the attempts each student makes in arriving at his or her solution to one problem from the assignment. By reviewing the process undertaken, one can see both the time spent on this problem as well as the missteps and misunderstandings along the way. In this example, the student spent roughly 30 minutes working on the formula, not counting what appears to be a break starting at 7:31 p.m.

Figure 5 shows the Aries log filtered sorted to identify similar submissions. Recall that Figure 3 showed Log # 16 and Log # 17 as being from the same Log Set. This means that the first several entries of the log had been automatically verified to be identical, signaling that one of the students shared his or her worksheet with the other student.



| | Log | Step | Time Stamp | Elapsed | Sheet Name | Cell Ad | Formula | Value |
|---|---|---|---|---|---|---|---|---|
| 1669 | 16 | 1 | 9/22/2013 22:28 | | Lookup Functions | F9 | =VLOOKUP(E9,$K$10:$M$15,2,TRUE) | 8 |
| 1670 | 17 | 1 | 9/22/2013 22:28 | | Lookup Functions | F9 | =VLOOKUP(E9,$K$10:$M$15,2,TRUE) | 8 |
| 1671 | 16 | 2 | 9/22/2013 22:29 | 0:00:32 | Lookup Functions | F10:F33 | =VLOOKUP(E10,$K$10:$M$15,2,TRUE) | 7 |
| 1672 | 17 | 2 | 9/22/2013 22:29 | 0:00:32 | Lookup Functions | F10:F33 | =VLOOKUP(E10,$K$10:$M$15,2,TRUE) | 7 |
| 1673 | 16 | 3 | 9/22/2013 22:30 | 0:01:02 | Lookup Functions | G9 | =VLOOKUP(E9,$K$10:$M$15,3,TRUE) | $400 |
| 1674 | 17 | 3 | 9/22/2013 22:30 | 0:01:02 | Lookup Functions | G9 | =VLOOKUP(E9,$K$10:$M$15,3,TRUE) | $400 |
| 1675 | 16 | 4 | 9/22/2013 22:30 | 0:00:09 | Lookup Functions | G10:G33 | =VLOOKUP(E10,$K$10:$M$15,3,TRUE) | $150 |
| 1676 | 17 | 4 | 9/22/2013 22:30 | 0:00:09 | Lookup Functions | G10:G33 | =VLOOKUP(E10,$K$10:$M$15,3,TRUE) | $150 |
| 1677 | 16 | 5 | 9/22/2013 22:31 | 0:00:30 | Lookup Functions | H9 | =HlooKUP | #NAME? |
| 1678 | 17 | 5 | 9/22/2013 22:31 | 0:00:30 | Lookup Functions | H9 | =HlooKUP | #NAME? |
| 1679 | 16 | 6 | 9/22/2013 22:31 | 0:00:40 | Lookup Functions | I9 | | |
| 1680 | 17 | 6 | 9/22/2013 22:31 | 0:00:40 | Lookup Functions | I9 | | |
| 1681 | 16 | 7 | 9/22/2013 22:32 | 0:00:24 | Lookup Functions | I9 | 400 | $400 |
| 1682 | 17 | 7 | 9/22/2013 22:32 | 0:00:24 | Lookup Functions | I9 | 400 | $400 |
| 1683 | 16 | 8 | 9/22/2013 22:33 | 0:01:08 | Lookup Functions | H9 | =HLOOKUP(C9,$O$9:$Q$18,D9,FALSE) | $800 |
| 1684 | 17 | 8 | 9/22/2013 22:33 | 0:01:08 | Lookup Functions | H9 | =HLOOKUP(C9,$O$9:$Q$18,D9,FALSE) | $800 |
| 1685 | 16 | 9 | 9/22/2013 22:33 | 0:00:19 | Lookup Functions | H10:H33 | =HLOOKUP(C10,$O$9:$Q$18,D10,FALSE) | $1,100 |
| 1686 | 17 | 9 | 9/22/2013 22:33 | 0:00:19 | Lookup Functions | H10:H33 | =HLOOKUP(C10,$O$9:$Q$18,D10,FALSE) | $1,100 |
| 1687 | 16 | 10 | 9/22/2013 22:34 | 0:00:39 | Lookup Functions | I9 | 1200 | $1,200 |
| 1688 | 17 | 10 | 9/22/2013 22:34 | 0:00:39 | Lookup Functions | I9 | 1200 | $1,200 |

Figure 5: Ares log filtered to show student collaboration

## 7. VALUE OF LOG DATA

One specific situation in the course is worth examining in detail. In this situation students are asked to use the AND function with three arguments. Each argument references a cell containing a Boolean value. As indicated in Figure 6, the AND function checks to see if each of the three arguments are TRUE or FALSE. Only if all the arguments are TRUE will the resulting value from the AND function be TRUE (see the Formula column).

While examining the log for this problem, researchers observed student attempts with quotes around the word TRUE. It is not uncommon for students to use quotes for any non-numeric text within a function. But, the use of quotes around TRUE indicates the student has a misunderstanding regarding data types. More specifically the difference between Strings and Boolean values. The word TRUE without quotes represents a Boolean value. Using quotes around the word TRUE changes the expression type from a Boolean to a String. Looking only at the submission data, 13 students' final attempts include quotes (") around the word TRUE. These students clearly have a misunderstanding regarding data type and submitted an incorrect solution to the problem. While the submission

ISSN: 2167-1435
v31 n3086

data indicates that only 13 students affected, the log data shows a different view of this misunderstanding. The log data identifies 44 students whose attempts include quotes around the word TRUE. This represents a much larger problem than the 13 students identified from the submission data based solely on the final solution. The duration of time these students spent solving the problem ranged from 3 minutes to over an hour.

| Log# | Step# | Time Stamp | Worksheet | Cell | Formula | Value |
|---|---|---|---|---|---|---|
| 72 | 9 | 2/4/14 8:11:38 PM | Boolean | K11 | AND(H11 | AND(H11 |
| 72 | 10 | 2/4/14 8:12:07 PM | Boolean | K11 | =AND(H11="TRUE",I11="TRUE",J11="TRUE") | FALSE |
| 72 | 16 | 2/4/14 8:26:46 PM | Boolean | K11 | =AND(H11="FALSE",I11="FALSE",J11="FALSE") | FALSE |
| 72 | 17 | 2/4/14 8:26:52 PM | Boolean | K11 | =AND(H11="TRUE",I11="TRUE",J11="TRUE") | FALSE |
| 72 | 23 | 2/4/14 8:33:01 PM | Boolean | K11 | =AND(H11="TRUE",I11="TRUE",J11="TRUE") | FALSE |
| 72 | 24 | 2/4/14 8:34:07 PM | Boolean | K11 | =AND(H11="TRUE",I11="TRUE",J11="TRUE") | FALSE |
| 72 | 25 | 2/4/14 8:34:18 PM | Boolean | K11 | =AND(H11="TRUE",I11="TRUE",J11="TRUE") | FALSE |
| 72 | 32 | 2/4/14 8:51:43 PM | Boolean | K11 | =AND(E11="TRUE",I11="TRUE") | FALSE |
| 72 | 34 | 2/4/14 8:56:16 PM | Boolean | K11 | =AND(E11="yes",I11="TRUE") | FALSE |
| 72 | 36 | 2/4/14 9:04:11 PM | Boolean | K11 | =AND(E11="yes") | FALSE |
| 72 | 41 | 2/4/14 9:18:01 PM | Boolean | K11 | =AND(H11="TRUE",I11="TRUE",J11="TRUE") | FALSE |
| 72 | 43 | 2/4/14 9:20:51 PM | Boolean | K11 | =AND(H11=TRUE,I11=TRUE,J11=TRUE) | FALSE |
| 180 | 31 | 2/18/14 11:14:14 PM | Boolean | K11 | =AND(H11="true",I11="true",J11="true") | FALSE |
| 180 | 33 | 2/18/14 11:16:21 PM | Boolean | K11 | =AND(H11="true",L11="true",J11="true") | FALSE |
| 180 | 35 | 2/18/14 11:17:15 PM | Boolean | K11 | =AND(H11="True",L11="True",J11="True") | FALSE |
| 268 | 59 | 2/16/14 5:06:03 PM | Boolean | K11 | =AND(H11="TRUE",I11="TRUE",J11="TRUE") | FALSE |
| 268 | 62 | 2/16/14 5:09:27 PM | Boolean | K11 | | |
| 268 | 63 | 2/16/14 5:10:37 PM | Boolean | K11 | =AND(H11="TRUE",I11="TRUE",J11="TRUE") | FALSE |
| 268 | 68 | 2/16/14 5:11:27 PM | Boolean | K11 | =AND(H11="True",I11="True",J11="True") | FALSE |
| 268 | 71 | 2/16/14 5:12:52 PM | Boolean | K11 | =AND(H11="TRUE",I11="TRUE",J11="TRUE") | FALSE |
| 268 | 72 | 2/16/14 5:12:53 PM | Boolean | K11 | | |
| 268 | 76 | 2/16/14 5:19:02 PM | Boolean | K11 | =AND(H11="TRUE") | FALSE |
| 268 | 78 | 2/16/14 5:19:21 PM | Boolean | K11 | =AND(H11=TRUE) | FALSE |
| 268 | 79 | 2/16/14 5:19:24 PM | Boolean | K11 | =AND(H11="TRUE") | FALSE |
| 268 | 81 | 2/16/14 5:19:32 PM | Boolean | K11 | =AND(H11=TRUE) | FALSE |
| 268 | 84 | 2/16/14 5:19:42 PM | Boolean | K11 | =AND(H11="TRUE") | FALSE |
| 268 | 85 | 2/16/14 5:20:11 PM | Boolean | K11 | =AND(H11=TRUE,I11=TRUE,J11=TRUE) | FALSE |
| 281 | 19 | 1/26/14 5:18:22 PM | Boolean | K11 | =AND(H11="TRUE",I11="TRUE",J11="TRUE") | FALSE |
| 281 | 21 | 1/26/14 5:18:50 PM | Boolean | K11 | =AND(H11="TRUE",I11="TRUE",J11="TRUE") | FALSE |
| 281 | 22 | 1/26/14 5:18:53 PM | Boolean | K11 | =AND(H11="TRUE",I11="TRUE",J11="TRUE") | FALSE |
| 281 | 24 | 1/26/14 5:19:47 PM | Boolean | K11 | =AND(H11="TRUE",I11="TRUE",J11="TRUE") | FALSE |
| 281 | 27 | 1/26/14 5:21:31 PM | Boolean | K11 | =AND(H11="TRUE",I11="TRUE",J11="TRUE") | FALSE |
| 281 | 30 | 1/26/14 5:22:39 PM | Boolean | K11 | =AND(H11=TRUE,I11=TRUE,J11=TRUE) | FALSE |

Figure 6: Evidence of student misunderstanding

Log data for this problem shows more instances of the misunderstanding than an analysis of the submission data might suggest. Implications of this finding can be important not only for identifying student errors but also for improving instruction or planning remedial actions. In this case the log data showed that the extent of this misunderstanding was greater than what might be expected based solely on the last solution attempts submitted. Instructional intervention can be designed and implemented to correct this misunderstanding not just for the 13 students whose final submission data indicated a misunderstanding, but also for the 31 other students who experienced this problem. This and other misunderstandings like it often can go undiagnosed by an instructor when the activity-trace data is unavailable or unexamined.

## 8. SUMMARY

By delivering student assignments in an Excel workbook augmented with a Helios log and then aggregating those logs using Ares allows for overview and analysis of the process students take in arriving at their solutions. The amount of data available is considerable. This example of 79 real student submission, comprises 8,599 log entries. A quick look at the number of step each student took to solve the problem provides an easy way to identify students who may be struggling with the content, deeper analysis can lead to insights about common errors and inform the professor to make pedagogical adjustments or to take remedial action for specific students. Additionally, very compelling evidence cam be collected indicating when students are sharing work rather than completing the learning individually.

Current versions of Helios and Ares are freely available for non-profit, educational and research from the lead author. If used for scholarly research, contact the lead author for current recommended citation.

## 9. REFERNCES

Aleven, V., McLaren, B M., Roll, O., & Koedinger, K. (2004). Toward Tutoring Help Seeking; Applying Cognitive Modeling to Meta-Cognitive Skills. Human-Computer Interaction Institute. Paper 156. Retrieved from http://repository.cmu.edu/hcii/156

Baepler, P. & Murdoch, C. J. (2010). Academic Analytics and Data Mining in Higher Education. *International Journal for the Scholarship of Teaching and Learning, 4* (2), Art 17, 1-9.

Behrens, J. T., Mislevy, R. J., DiCerbo, K. E., & Levy, R. (2011). An evidence centered design for learning and assessment in the digital world. In M. C. Mayrath, J. Clarke-Midura, & D. Robinson (Eds.), *Technology-based assessments for 21st century skills: Theoretical and practical implications from modern research* (pp. 13–54). Charlotte, NC: Information Age.

Chung, G.K.W.K. (2014). Toward the relational management of educational measurement data. *Teachers College Record, 116*(11).

Johnson, L., Adams, S., & Cummins, M. (2012). Horizon Report: 2012 Higher Education Edition. Austin, TX: New Media Consortium.

Martinez-Maldonado, R., Clayphan, A., Yacef, K., & Kay, J. (2014). Towards Providing Notifications to Enhance Teacher's Awareness in the Classroom. In *Intelligent Tutoring Systems* (pp. 510-515). Springer International Publishing.

Sewell, A. (2002). Constructivism and Student Misconceptions: Why every teacher needs to know about them. *Australian Science Teachers' Journal. 48*(4), 24-29.

Siemens, G. (2012). Learning Analytics: Envisioning a research discipline and a domain of practice. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge - LAK '12*. ACM press, New York, NY, 4-8.

Siler, S. A., & VanLehn, K. (2014). Investigating Microadaptation in One-to-One Human Tutoring. *The Journal of Experimental Education*, (ahead-of-print), 1-24.

VanLehn, K. (2011) The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems. *Educational Psychologist, 46*(4), 197-221.

ISSN: 2167-1435
v31 n3086

©2014 EDSIG (Education Special Interest Group of the AITP), www.aitp-edsig.org
and FITE (Foundation for Information Technology Education), edfoundation.org/