Investigations in Mathematics Learning
©The Research Council on Mathematics Learning
Fall Edition 2015, Volume 8, Number 1

Integration through a Card-Sort Activity

Kris Green
St. John Fisher College
kgreen@sjfc.edu

Bernard P. Ricca
St. John Fisher College

Abstract

Learning to compute integrals via the various techniques of integration (e.g., integration by parts, partial fractions, etc.) is difficult for many students. Here, we look at how students in a college level Calculus II course develop the ability to categorize integrals and the difficulties they encounter using a card sort-resort activity. Analysis of the data required the use of several non-standard techniques which provided interesting insights into the ways students develop categories in mathematics. One finding of note is that students may need a significant amount of time "off topic" to allow sufficient time to fully organize their schema for integration.

Keywords: integration, categorization, card sort task, student learning

Understanding Students' Difficulties with Integration Through a Card Sort Task

Most second semester college calculus courses involve two main topics. The first part of the semester is devoted to techniques and applications of integration, while the second half of the semester is often devoted to topics related to sequences and series. While the need for this particular content will not be debated here, it is sufficient to point out that that (a) not all college calculus courses work this way (e.g. Heid, 1988) and (b) not everyone is convinced that it is necessary to devote that much time to these topics in the days of Maple, Mathematica, Wolfram-Alpha, and other Computer Algebra System [CAS] tools (Gordon, 1993).

Perhaps, in a CAS-driven world, the need for teaching procedural skills such as integration should be focused on particular groups of students who are expected to need these skills in the future, rather than all students (Gordon, 1993; Oates, 2009). As Harel (2008) points out, though, we might lose certain other aspects that are not directly related to the knowledge of integration. And we might not be aware of which students will need what skills in the future, making such tracking difficult.

One difficulty faced by students when asked to integrate a function is in selecting an appropriate technique to apply. Typical techniques taught in a second semester calculus class include integration by substitution, integration by parts, partial fractions, trigonometric substitution, and the like. In a brief survey of available calculus books at the present, almost all present the various techniques of integration, then include a discussion of how to organize all of these into a heuristic for dealing with integrals in general. These are typically based on some version of Shoenfeld's (1978) or Slagle's (1963) algorithm.

However, when teaching these algorithms, students suffer from several pitfalls which many cannot overcome. First, if a student's algebra skills are weak, implementing the methods of integration, once one is selected, becomes very difficult and errors in basic algebra or differentiation can mislead them. Asking students to check their answers by differentiation then requires them to rely on this already weak skill and rarely helps them to find their mistakes or correct them. Further, a given integral can be expressed in multiple forms. Students with weak algebra skills can have difficulty recognizing how two answers that appear different may be the same, so "looking in the back of the book" often fails to help them identify the problem.

Another concern is more conceptual in nature and calls to question the entire undertaking. By teaching these explicit techniques of integration, structured in some general algorithm, we can, at best, hope to turn our students into slower, incomplete versions of a CAS that are prone to basic computational errors. Given the extensive time devoted to these topics, we should consider whether this time is well spent, especially in light of the general audience for these courses.

Based partly on these concerns and partly on curiosity about several inquiry-based labs we encountered for teaching techniques of integration (e.g. Anderson, 2010; Stewart, 2008), the present study examines how students develop understanding of techniques for integrating functions through two unique components. First, we restructured the typical approach to the content to build a scaffold for the entire concept of "integrating functions." This structure was then filled in by focusing on the different components of this process. Second, the scaffolding gave us access to a large variety of data on students' concepts of integration that required non-standard tools for analysis. We thus applied techniques adapted from graph theory (Green & Ricca, 2014) to examine how these students developed the ability to classify integrals by the method for solving them.

In particular, we focused first on classifying integrals before explicitly teaching the specific techniques for solving each type of integral problem. On the surface, this may seem illogical, but there are two factors that led us to explore this approach. First, the availability of the CAS provides a tool for students to collect data, even when they cannot themselves perform the integration. They can then construct a classification scheme. Many CAS have a "verbose mode" available whereby the steps taken by the CAS can be made explicit, walking the user through the solution, providing additional data. Second, by creating the classification scheme first, we seek to unify the different techniques of integration into a single coherent framework from the start, rather than try to reorganize the way students think about integration after they have learned specific details. By scaffolding the process, we expect that students will develop a general scheme that motivates the need for different techniques of integration, so that when these are encountered later, they can be used more efficiently.

We thus had two research questions to investigate. How does a "classification first focus" – rather than learning specific techniques first – impact students' abilities to classify integrals by the method of integration? How do students organize their knowledge of integrals to aid them in solving routine integration problems? These questions were explored by gathering data through an inquiry activity based on cycles of a card sort task (CST) with indirect feedback.

Background

Recently, Sofronas, DeFranco, Vinsonhaler, Gorgievski, Schroeder and Hamelin (2011) found agreement among most national experts that integration is a critical component of learning what the calculus is all about. Specifically, these experts ranked techniques of integration and concepts of integration as the two most important aspects of learning integration. However, previous studies have shown that students' conceptual understanding of integration is weak (Orton, 1983; Bennett, Moore, & Nguyen, 2011) We know that a focus on procedural knowledge is inadequate without a concomitant development of conceptual and strategic knowledge for helping students decide what algorithms and techniques to use in particular situations (Shoenfeld, 1978; Rittle-Johnson & Alibali, 1999). Although procedural and conceptual knowledge interact dynamically (Byrnes & Wasik, 1991) without parallel knowledge structures in procedural techniques and strategic thinking, students will not be developmentally ready to use their procedural knowledge (Adey, 1999; Pettersson, & Scheja, 2008; Campbell, 2011). Acceptance of this has led to the development of and teaching of more transferable algorithms for teaching the strategic aspect of integration. However, many of these can devolve into purely procedural approaches or random guessing. Even when students are good at implementing the algorithm, external features in presenting a problem can mislead them (Lauten, Graham, & Ferrini-Mundy, 1994; Mallet, 2011) especially in the absence of adequate strategic knowledge. This then forces the students to compartmentalize the various aspects of the knowledge of integration, leading to future problems and a lack of flexibility in problem solving (Mallet, 2011; Tall, 1993; Abdul-Rahman, 2005; Rosken & Rolka, 2007). Even when students possess adequate procedural and strategic knowledge related to integration, fundamental difficulties with algebra interfere with their implementation (Grundmeier, Hansen, & Sousa, 2006).

Thus, while some studies call for more development of procedural and strategic knowledge (e.g, Grandsard, 1997), this is largely the opposite of creative problem solving. We are attempting to reduce an art – solving a particular integration problem – to an algorithm (what Buschberger (1990) calls "trivialization"). We cannot expect, however, that students will develop sufficient fluency with this to rival the Risch algorithm implemented by most CAS, nor should we expect students to merely reproduce the complexities of this algorithm. If we expect students to exhibit strategic thinking about how to solve a problem, their existing strategies tell them to use a better tool – like a CAS – rather than rely on their own inadequate abilities. Instead of forcing them to focus on the procedural side of things, we must recognize that even a procedural topic like integration requires creativity to solve. Consider, for instance, the case of integrating $\sin^2 x$. This requires much more than a list of procedures to follow, even though it can be reduced to such a list. Recognition that substitution of an equivalent expression for the integrand will "simplify" the problem – a term itself which is confusing to students since the integrand grows from one term to two terms – is not a trivial thing to expect. Trying to memorize all the patterns without strategic thinking about their uses to help organize them will result in overload and failure. After all, Gradshteyn & Ryzhik (1994) is not a book to be memorized!

One of the problems in much of the research on students' knowledge of integration is the conflation of two things that are not necessarily equivalent. Studies like Grundmeier, Hansen, & Sousa (2006) and Mahir (2009) observe student work on a sample of integration-related procedural questions then draw conclusions about the students' underlying conceptual understanding. We certainly expect these inferences to be valid in many cases, but we recognize that a single visible symptom can be caused by several different underlying misconceptions or mistakes. In the GLP metaphor for learning (Green & Ricca, 2013), we refer to this as the confusion between the underlying space of possibilities (genotypes) and observed behaviors (phenotypes). So, in order to help students resolve their confusion and construct more appropriate conceptual, procedural and strategic thinking about a topic like integration, we need to understand their current thinking more deeply. We must find out about the underlying "genes" from which they have built their knowledge and skills. We must support the evolution of their existing set of genes into a more fit and successful phenotype not only by helping them to acquire new genes for thinking, but also by selectively mutating and modifying the existing thinking.

But how do we guide students in developing problem-solving skills relevant to integration? First, we must make sure that we level the playing field so that all students are targeting the same goals in problem solving. This means that we must help them see that integration is a tool for solving other problems and that conceptual knowledge of integration (e.g. as a limit of sums) will help set up such problems while procedural and strategic knowledge will help solve the problems. Without explicitly setting such goals, subjectively perceived external factors, such as efficiency and accuracy, may drive the problem-solving process (Campbell, 2011; Pettersson & Scheja, 2008). Second, we must help them develop all three legs of support for their knowledge of integration – procedural, conceptual and strategic – to avoid a wobbly foundation and ensure that all students have sufficient potential (Pettersson & Scheja, 2008) to develop these understandings. This means a focus on the deep knowledge of integration and building a coherent understanding of integration as a single concept rather than separate ideas that conflict (Tall,1992). Such deep learning focuses on the conceptual and strategic support for the topic, rather than strictly on the procedural methods (Mahir, 2009; Mackie, 2002).

What would a calculus classroom driven by such ideas look like? Several studies (Palmiter, 1991; Heid, 1988; Murphy, 1999) have reported on experiments that integrate CAS support for procedural skills while focusing primarily on the problem solving and conceptual development of the students. Only after a majority of the course dealt with conceptual understanding and application were the students exposed to the procedural techniques used in integration by hand. Findings suggest that the CAS does not reduce the amount of procedural knowledge gained, in spite of the shorter time spent on this aspect. In particular, students in Heid (1988) recognized the CAS let them focus more on problem solving, which is often cited as one of the primary goals of mathematics education. While studies such as these are promising, others (e.g., Meel, 1998) have shown more mixed results when implementing a CAS-enhanced calculus curriculum. In the end, though, we find few studies investigating the co-development of procedural and strategic knowledge in integration. Most focus on conceptual understanding, following Orton (1983), and assuming implicitly either that we are already doing the best we can with respect to the procedural side of integration, or that understanding this aspect of student learning is too difficult to tackle.

Functional MRI studies by Schroeder (2011) have identified several areas of the brain active in the integration tasks after explicitly training the participants in Shoenfeld's (1978) strategies for integration; this training in strategic thinking also led to improvement in participants' accuracy. Further studies have shown that in some cases problem-solving skills can be improved by developing supporting conceptual knowledge. This is likely due to an improved problem representation (Rittle-Johnson & Koedinger, 2005). Others have shown that strategic knowledge can alter problem representation (Alibali, Phillips, & Fischer, 2009) although none have explored this development at the level of the calculus.

One way to elicit students' knowledge of categories, such as the categories of different integral problems (simple, substitution, integration-by-parts, etc.), is to use a cart sort task. In such tasks, participants are given a set of cards, pictures, or objects and asked to sort them on some criterion. Often, they are asked to repeat the sort, using a different criterion each time, although Rugg and McGeorge (1997) mention a number of variations on this basic task. Once the participants have sorted the cards into groups, researchers have a number of tools available to make sense of the different sorts. Fincher and Tenenberg (2005) review a number of these in their introduction to a special issue of *Expert Systems*; these largely fall into semantic or syntactic analyses. Often, techniques from graph theory, such as edit distance (Deibel, Anderson & Anderson, 2005) to compare the participant sorts either on a participant-by-participant basis, a sort-by-sort basis, or in comparison to some "expert" sort of the cards. Since we wanted to focus on participants conceptual knowledge categories to support their strategic knowledge of integration techniques, we implemented a "free sort" (Harper, Jentsch, Berry, Lau, Bowers, & Salas, 2003) with indirect feedback over multiple rounds of sorting.

Methodology

Participants

Participants in this study were students enrolled in a second-semester calculus course at a small, liberal arts college. The course was taught by one of the authors. The second author has previously taught this course at the college, and served as an observer during the activities described below. All students were provided an opportunity to opt out of the research experience, which was conducted as a normal part of the course activities and approved by our institution's IRB procedures. None of the fifteen students opted out, providing a sample of 15 students. Their previous calculus backgrounds and intended majors are shown in table 1. Note that the majority of the participants had no previous experience with calculus II. Most of the participants (8 of 15) planned to major in chemistry, which is quite typical of students enrolled in calculus at our institution.

Data Collection

The activities described below took place over a period of 15 class meetings and were supplemented with work outside of class in the form of pencil-and-paper homework and WebAssign problems. The majority of the in-class activities were conducted in a computer lab using Maple. The class met four times each week with 55 minute class meetings. The timeline and support for each component of the activity, along with a brief description are provided in table 2. In all components, students were given the option to work with a partner or alone.

Table 1. *Summary of participants' experience with calculus*

Credit for calculus I from…	Never taken calculus II	Had taken calculus II
AP exam	3	0
Calculus I at our college	5	3
College course taken in high school	2	1
Calculus I at another college	1	0
TOTAL	11	4

Prior to this, students had studied chapters 5 and 6 of Stewart's Calculus book (2008). This gave them familiarity with the basic definition and concept of integration, integration of basic functions, some familiarity with substitution, several applications of integration, and the fundamental theorem of calculus. However, they had not studied or practiced integration by parts or the other specialized techniques of integration found in chapter 7 of the textbook.

After a short pre-test on differentiation and integration to ensure that participants had an adequate understanding of the symbols involved in integration and differentiation, which is a pre-requisite for using card sort tasks (Rugg & McGeorge, 1997), participants completed the demographic survey. They then began round 1 of the experiment. For this, the students received twenty cards, each with an integral printed on it; these are reproduced in the appendix. Each card had been assigned a random three-digit number to facilitate identification. Students also received a worksheet with directions and room to write down the results of their sort activity with explanations (see Table 3 for the directions given at each phase.) The integrals on the cards were chosen from five categories, based on the primary technique of integration needed: simple integrals, integration by simple substitution, integration by parts, integration by partial fractions, and those requiring a mix of methods. In addition, the integration variable was randomly chosen to be either x, r, t, or v in order to evaluate whether surface features dictated some classification schemes. Each of the subsequent four rounds began with a data collection activity – either using the CAS to get information about each integral or doing some work by hand on each integral – and then revisiting their classification scheme to account for the new data.

Table 2. *Timeline of the data collection and classroom activities involved in the project. Days are indexed from the first day of class.*

Day	Setting	Description	Data Generated
12	Assignment	Pre-test on differentiation and integration.	Demographic questionnaire Pre-test on differentiation
13	Computer lab	Initial card sort activity [15 min] Second sort activity, using Maple to find each integral [30 min]	Round 1 sort and explanation Round 2 sort and explanation Observational data
14, 17	Classroom completed at home	Third sort activity, taking derivatives of each answer from phase 2	Round 3 sort and explanation Observational data
18	Computer lab	Fourth sort activity, using the "verbose" mode of Maple to see the steps in the integration	Round 4 sort and explanation Observational data
19-26	Classroom	Instruction and practice on techniques and applications of integration	Various homework assignments and class activities
27	Classroom	Final sort activity	Round 5 sort and explanation Observational data
28	Classroom	Unit test on integration	Test 2 results
50	Classroom	Final exam	Performance on four exam questions

Note. Days 15 and 16 involved a review and exam for the first unit of the course which did not include content from the activities discussed in the present paper.

After the first four rounds were completed, the instructional period consisted of "active lectures" on the various techniques of integration along with practice applying each technique. During this period, the students submitted work on various integration problems collected in notebook in which they were instructed to include all their work; if they started off and realized they were on the wrong track, they were instructed to draw a single line through the work that was incorrect and start over, carefully labeling the problems. After the instructional period, but before the exam on this material, students completed the final round of sorting.

Table 3. *Instructions provided at each round of the card sort activity.*

Round	Directions
1	There are lots of methods for integrating functions. Look at the 20 integrals on the cards provided. Your job is to use whatever tools and techniques you want to try and group these integrals into categories so that the integrals within each category require a similar method to carry out the integration.
2	Now use the CAS to find the antiderivative of each function. Then revise your groups from part 1, using the antiderivatives provided by the CAS to guide you. Explain the reasoning behind your new groupings.
3	Next, take the derivatives of each of the antiderivatives you computed in part 2. Take note of any special techniques used to find the derivatives. (Hint: In each case, you should be able to recover the original integrand by taking the derivative…) Revise your groups from part 2, using your notes from taking the derivatives.

4	Finally, you are going to have the CAS show you all the steps used in determining the antiderivative of each function. This can be accomplished with the script below. Then, revise your categories and explain your new groupings for the integrals.
5	Group the following integrals into categories by putting an "X" in the box corresponding to the category of integrals that require the most similar techniques for solving. Then, describe each of your categories in the space below. You do not have to use all six categories; they are provided as a convenience.

Results and Analysis

After collecting the data described above, we recognized that we had collected a tremendous amount of data, but that many of the standard analysis techniques either did not apply or resulted in a loss of resolution. For example, using only edit distance (Deibel, et al, 2005) does not let us analyze a student's groupings for both incorrect and correct groupings, since we only determine the number of moves required to match another grouping. The analysis of many card sort tasks focuses on the number of correct groupings only, which then ignores the information about student thinking and learning encoded in the incorrect groupings.

Primarily, we needed ways to visualize the groupings of each student at each phase, and of the entire class in a variety of ways. Our analysis included standard methods, such as the edit distance (Deibel, et al, 2005) as well as a variety of non-standard methods. These are briefly discussed here, and more fully discussed in Green & Ricca (in process). The methods are adapted from various graph theory and network theory tools as well as non-standard tools, including the construction of link-gaps graphs, minimum volume enclosing ellipsoids, affinity networks, cluster analysis, and Monte Carlo methods.

The non-standard methods were implemented to address shortcomings in the standard approaches with respect to what we wanted to explore. Simply looking at the number of integrals that each student classified correctly ignores a great deal of the information available since students could construct their own categories that might differ in number from each other, from the ideal grouping, and from themselves between rounds. Thus we wanted a method of analysis that allowed us to see this information. In our activity, there is an ideal grouping – one consisting of five groups of four integrals each – so we could use edit distances (Deibel, et al, 2005) or the Jaccard index (Real, 1999) to compare each student's adjacency matrix with the ideal. However, these measures conflate conceptually independent aspects of this task that we wanted to study separately, in that they provide only a single number for the "distance" between matrices like our data. This does not distinguish, for example, whether a student's score improved by getting more groupings correct or fewer incorrect groupings. As we shall see, methods that allow one to examine such distinctions provide insight into student understanding.

One of the primary graphical tools used was the use of a link-gaps graph. For each of the 75 student groupings, we compute the number of correct links present; each link represents a pair of integrals that the student has placed in the same group. We also compute the number of correct gaps present – that is, we count the number of integrals that are correctly not grouped together. Representing these as a fraction of the possible correct links (30) and gaps (160) we then plot each student's grouping at each round as a point on a Cartesian graph. The point (1, 1) represents the ideal grouping. Such a graph allows us to explore how the students clustered from round to round to investigate how feedback and instruction impacted the classifications. Details of this approach are available in Green & Ricca (2014).

Some reference points may help to understand the plots better. Suppose a student constructed a grouping that exactly matched the ideal grouping. That student would be plotted at (1, 1) on the grid. If a student put each integral in a separate singleton group, they would be located at (0, 1). If a student put all the integrals into a single large group, they would be plotted at (1, 0). Getting four groups right but splitting the last group into two pairings would result in a point at (0.9, 1), while placing one from each correct group into a new group to make four groups of five that are not related in the ideal grouping would plot at (0.25, 0.75). Figure 1 shows the basic links-gaps graph with data from the five rounds of the experiment.

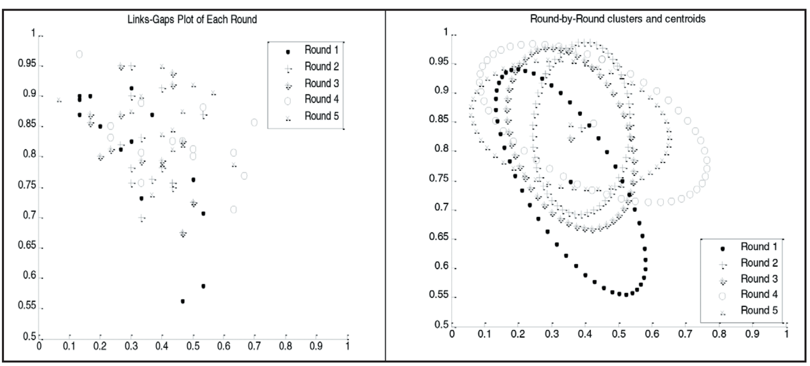

Figure 1. Links-gaps plot of student scores (X = % of correct links, Y = % of correct gaps) over each round, treating singleton groups as incorrect. [Left] Shows all students at each round; [Right] shows minimum volume enclosing ellipses with centroids for each round.

Clustering of the students was explored by constructing the minimum volume/area enclosing ellipses using Moshtagh's (2005) algorithm. This provided information about the size (via the minor and major axes), location (via the centroid), and orientation of each cluster. The centroids provide a measure of typical student results, allowing us to explore how performance changed during the different rounds of the activity, taking the class as a whole entity.

We also use the area of each ellipse, computed as a percentage of the total links-gaps graph area, as a measure of the variability of the student groupings at each round. The raw data and the round-by-round clusters are shown in figure 1, which shows that overall, the student groupings do not appear to change much over the five rounds of the activity, but this is misleading.

Table 4. *Summary data of measures at each round.*

Round	Centroid (N = 15)	Area of Cluster (%)	Average Edit Distance to Ideal	Average Edit Distance to Round	Average Number of Groups
1	(0.52, 0.73)	9.47	10.40 (1.24)		5.1 (1.13)
2	(0.55, 0.82)	6.71	10.00 (1.89)	7.73 (3.20)	4.7 (0.88)
3	(0.49, 0.82)	8.77	9.47 (1.68)	9.67 (1.88)	5.1 (1.46)
4	(0.63, 0.80)	12.24	8.67 (1.91)	10.00 (1.60)	4.9 (0.96)
5	(0.45, 0.84)	10.06	8.20 (1.70)	9.40 (1.12)	5.6 (0.91)

Note. Area is shown as a percentage of the total links-gaps plot area. For edit distance, standard deviation is shown in parentheses.

The centroids and variability measures are shown in table 4. Note that the areas of the clusters are fairly consistent, around 10% of the total area, throughout the five rounds of the activity. While the edit distances do decrease slightly from the beginning to the end, the spread of these remains approximately constant, and one can see that the edit distances between successive groupings are about the same size as the edit distances to the ideal. One difficulty in interpreting these clusters is to decide whether the movement of a student through the five rounds of the experiment is "significant" and what significance actually means. This interpretation is still in process (see Ricca & Green, under review), but relies on a Monte Carlo simulation to explore the concept-space mapped by the links-gaps graphs in order to understand how much movement one would expect at random in comparison with the movement observed. Another complication is the existence of an additional degree of freedom in the student work: the number of groups a student used could change and affect the interpretation of the data. Table 4 shows the number of groups at each round. Notice that the number of groups remains fairly stable, but the variation is much larger in rounds 1 and 3.

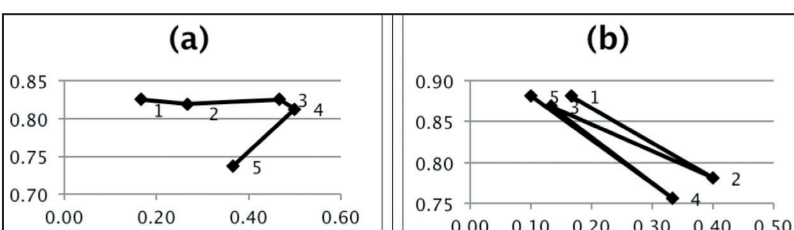

Figure 2. Links-gaps plot showing round-by-round trajectories of (a) a high-achieving student and (b) low-achieving student.

Some of this is apparent in following the trajectories of individual students. Figure 2 shows the round-by-round progress of two students on a links-gaps graph. The plot on the left are for a student who did relatively well on exam content related to classifying integrals and carrying out the integration; the data on the right show the path of a student who performed poorly on these tasks. Recall that the ideal grouping is in the upper right corner, at (1, 1), so both of these students exhibit changes that actually move away from the ideal. But they do so in quite different ways. Student (a) moves to the right (more correct groupings) and then starts to get more incorrect gaps, moving down. The student on the right almost oscillates; in between some rounds, the student improves the number of correct links while getting more incorrect gaps (moving to the right and down). In between other rounds, the student loses some correct links and also eliminates some incorrect gaps (moving up and to the left). Overall, this second student moved down and the left, which is away from ideal, while the first student has made some positive growth. This difference was borne out in the exam materials.

We constructed two empirical models for predicting student work on integrations problems on the unit test and the final exam. Each of these exams included four procedural problems and student totals on these were scaled between 0 and 1. Despite the small sample size ($N = 14$) each model had high predictive power. The models are summarized in table 6. In both models, a single outlier was identified and removed from the data. The variables in the models suggest a great deal about student learning in this context. That the number of correct links after round 3 was significant for both models is encouraging. These links were constructed by students after explicitly connecting a more familiar and easily implemented topic, differentiation, with a newer content, integration. We infer that students were able to make use of these connections to support their learning of the concepts, procedures, and strategies of integration.

The second predictor for the unit test was based on the number of correct links students constructed after receiving feedback from the computer on what the result of each integration would be. In some of these cases, the surface features to which they attended correspond with the deeper structures needed to understand how to integrate these functions. In the model predicting final exam performance, the number of correct gaps after the initial sort (round 1) was significant, suggesting that students with fewer pre-conceived notions about what makes two integrals similar allowed them more flexibility to modify their thinking throughout the unit of study. In a sense, these students had less to "unlearn."

We note that student scores from test 2 were not significant in predicting final exam scores on the related questions. We also note that students did quite well on the four procedural final exam questions when compared to performance on similar questions on the test immediately proximal to the

Table 6. *Summary of regression models predicting scores on the integration procedural problems from the unit test and final exam.*

Model	Range	Mean (Standard Deviation)	R^2	Variables
Unit test	0.325 – 0.750	0.571 (0.118)	0.7634	Links at round 3**
				Links at round 2**
Final exam	0.375 – 0.975	0.708 (0.194)	0.7891	Links at round 3***
				Gaps at round 1**

** $p < 0.01$, *** $p < 0.001$

unit of study (a mean of 0.708 versus 0.571, $p < 0.05$). We will have more to say on this later.

Table 7 reports detailed results from the four procedural questions on the final exam. For all four problems, we note that the majority of the students were able to both categorize the integral by the correct method needed to compute it as well as set up the integration technique needed correctly. Eight of the fifteen students were able to categorize all four of the problems correctly, while eleven of the fifteen correctly categorized at least three of the problems (see table 8).

Table 7. *Analysis of work on final exam questions showing the number of students in each performance group for each question.*

Success by Parts	Integration	Simplify	Simple Substitution	Trig Substitution
Useful category, implementation, algebra	9	7	4	2
Useful category, implementation	0	4	10	7
Useful category	0	0	0	5
Unhelpful category	6	4	1	1
Average Score (out of 10)	6.67	7.47	8.20	6.00

Note. "Useful category" refers to those students who categorization of the integral by type would allow for successful computation of the integral. "Implementation" refers both to the correct use of the method and selection of any options (such as the u and dv in IBP). "Algebra" refers to whether correctly performed all algebra, including differentiation for substitution, etc.

Table 8: Number of students correctly categorizing the four integrals on the final exam.

Problems categorized correctly	All four	Three	Two	One
Count of participants (N = 15)	8	3	3	1

Affinity graphs also provided insight into student thinking. These graphs are constructed by considering not only which integrals students group together in a particular round, but also how they label the groups. For example, figure 3 shows the affinity graph for round 4 of the study. The integrals are labeled by the group they belong to in the ideal grouping. Integrals in group "a" are simple integrals, those in "b" require simple substitution, those in "c" require integration by parts, those in "d" require partial fractions, and integrals in category "e" require more than one strategy. After coding students' category labels, we find groups with the labels shown above and below the integrals:

those categories listed above match the instructor "ideal" groups. The lines indicate the categories into which each integral was placed, and line thickness illustrates the proportion of students making that connection. Thus, by round 4, we see that many of the students have started to construct the categories that are "useful" and have also correctly placed many of the integrals into these categories. At the same time, we see confusion about category "e" and a few unhelpful ways of thinking that are persistent and focus on surface similarities in the integrals, rather than deeper structure.

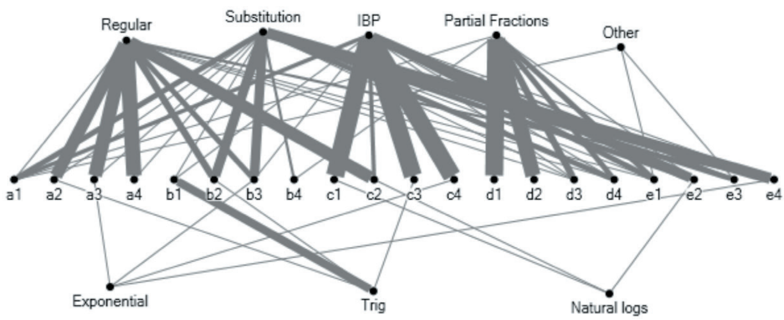

Figure 3. Affinity graphs showing each integral and the descriptions of the categories at round 4. The categories along the top row are the instructor labels; those in the bottom row are student category labels. (The edge thickness between an integral and its category is proportional to the number of students who placed that integral into the associated group.)

Discussion and Implications

Based on the edit distance (table 4) it seems clear that students did move closer to the ideal categorization throughout the activity. This movement seems to be due primarily to students making fewer incorrect pairings, rather than making more correct pairings. At the same time, it is clear that student classifications changed considerably from round to round. In many cases, students moved away from the ideal between rounds, while overall moving toward the correct grouping. Taken with the spread of the data at each round, we believe this variability means that the categories are not solid, which would suggest slower performance on tasks requiring this knowledge (Barasalou, 1983), but we did not collect data on the time it took students to perform any of the tasks described herein. It also appears that students are "resetting" between rounds and re-sorting based on some primitive ideas (p-prims) rather than incrementally modifying their existing framework. These underlying ideas may be related to the p-prims of diSessa (2001). For example, one group of students began round 2 with their cards in order from round 1 before adjusting for the new information; yet both of these students' groupings at round 2 were a large edit distance from their round 1 sorts, showing almost half of the cards changed groups. Such variations suggest that variability is underestimated in most studies, which could explain why applying educational research to particular settings does not always match expectations.

The variability and movement from round-to-round also demonstrates that learning is not monotonic. Students rarely move toward expected understandings in a uniform manner. This is likely due to a great many factors, such as individual confusion and misconceptions. Generic classification skills (if they exist) could also be playing a role. Adey (1999) demonstrated that the development of formal reasoning skills in concert with content results in commensurate gains for both aspects of thinking. We expect that another aspect of all this is that students are working with ideas at a deeper level than what is visible, a level sometimes referred to as p-prims (diSessa, 2001). Thus, sometimes, students are adjusting their category definitions, based on the underlying structure, and sometimes they are refining the membership of the categories. Certainly both are related; we note other examples of students acting consistently with this interpretation, (e.g. McNeil et al. 2010). They may not be very good at checking consistency with their own categories; this could be due partly to the nature of categories that are created and used "on the fly" (Baralou, 1983) or the difficulty with seeing category definitions as both a way of *including* certain items and *excluding* others (e.g., Edwards & Ward, 2004).

From the affinity graphs (see figure 3 for an example) we note that students' categorizations exhibited three features to differing degrees. Some are based on what we deem to be salient features, which could support student strategic knowledge for implementing integration. Still other categorizations seem based on surface features that are not appropriate for carrying out integration, such as using the functions involved, rather than the algebraic structure of the integrand. These categories are not entirely unexpected based on their prior mathematical knowledge. But we also note a third set of classifications that we refer to as sympathetic, since closer examination suggests reasons for the incorrect classification of some. For example, the integral of the natural logarithm requires integration-by-parts to implement; however, this integral is often listed in the "basic integrals" on the inside cover of many calculus texts. So we expect that some students have simply memorized the result of the integral, rather than the process of integrating this function. These three categories seem closely linked to the idea of heaps, complexes, and concepts (Vygotsky, 1986)

By the end of the calculus course investigated in this study, we note that students were able to categorize and implement correct methods for integration, but not immediately after the instructional period – some time was needed for them to organize and modify their schema. More study of this time period is needed to investigate changes in student thinking. We attempted to predict student performance on the final exam based on their work during the study and their performance on the unit exam. Such predictions were terribly inaccurate, possibly because we were trying to predict an evolutionary path of learning, possibly because there is a need for adequate wait time to allow the students' brains to reorganize the concepts and procedures and these reorganizations are not linearly related to students' precursor learning. Of note, is that actual final exam scores were only significantly related to the number of correct pairings at round 3 and the number of correct gaps made at round 1; both had impacts of a similar magnitude (see table 5). The first factor includes all the components needed: surface features, results of integration, and the related derivative rules; the second factor includes what they have to overcome or unlearn. Thus, it helps students to have more gaps correct at the initial categorization. We believe this is related to having fewer pre-conceptions about the problem with more openness to changing one's categories demonstrating flexible thinking.

An interesting aside is that, in spite of working in pairs, no pair of students constructed identical groupings across the 20 cards for any of the five phases of the CST. Some did share individual groups, but none had all groupings identical to another student on any round. Thus, despite sharing ideas, each student developed his or her own ideas. Thus, group interactions seem to have resulted in learning, but the mechanisms and reasons are poorly understood.

Increasing student ability to choose a technique of integration and correctly implement it on a given integration problem is not simply the accretion of more and more techniques along with a heuristic to choose among them. Instead, there are underlying components that must be appropriately coordinated to be successful at integration. Thelen and Smith (1994) undertook a close examination of how various forms of human locomotion (e.g., walking, running, jumping) develop. Their examination highlights the dependence of all these forms on the contextual coordination of an underlying set of muscle movements. While there are broadly applicable sequences of development (e.g., walking occurs before running in almost every person), Thelen and Smith showed that one form of locomotion is not built from a previously existing ones. Instead, each form of locomotion coordinates the underlying muscle movements in new ways and the basis of each coordination is contextual, dependent upon what the person wants to do, and what environmental constraints are in place. Thelen and Smith posit that such contextual coordinations are the basis of all cognition and action. Thus, although problems *a1* and *d1* from our integration tasks may appear similar on the surface (both are rational functions), direct polynomial division leads to two simple integrands in problem *a1*, while direct polynomial division is of little help in simplifying problem *d1*. Not only does this have implications for learning, it implies that useful formative assessments must examine not an hypothesized prior understanding upon which curriculum can be built, but instead, formative assessments must examine the underlying components that are to be coordinated.

We would be remiss to ignore certain limitations in the study, in design, analysis, and interpretation. With respect to the design of the activities, we note that feedback to the students was only indirect. We did not even implement a "you have X wrong" type of feedback. This was out of concern that students would focus too much on getting the right answer and not enough on developing their own conceptual understanding, but it might have been useful for helping them measure their progress.

Note that many card sort tasks include instructions that students first read each card separately, then start classifying. We did not include such instructions, giving the students more freedom in how they chose to implement the sort procedure. For example, some students executed a "sequential sort" by putting the first card in a group, then testing whether the second card belonged to that group. If it did not, a new group was formed for the second card. This was repeated with each subsequent card, adding new groups as needed. Other students began by looking at all the cards, then attempting to construct groups holistically. This lack of unity may have introduced too much of a confounding variable.

Also, we seem to be using card sorts in a non-standard way. Rather than attempting to elicit how participants group them based on techniques they would use from their prior learning (Rugg & McGeorge,1997) we are attempting to build their learning by categorization through scaffolding and feedback. Moreover, we are looking at aggregate data, which is more commonly done in closed sorts (Fincher & Tenenberg, 2005) than in open sorts like this one.

We also note that the different features students used for categorizing the integrals demonstrates the role that mathematical knowledge prior to integration (e.g. algebra skills) may actually have hindered them by focusing them on surface, rather than salient features. Others have noted similar situations, where previous learning activates earlier, less helpful conceptions of mathematical structures (McNeil, Rittle-Johnson, Hattikudur, & Petersen, 2010). Interpretation is further complicated by the awareness that, for certain integral problems, surface features become more salient. For example, integrands which are rational functions typically require integration by partial fractions. Thus, there are complex interactions between the form of a problem and the method used to solve the integration problems.

Students described their categories in their own words. While this freedom provided insight into student thinking, some descriptions required interpretation by the researchers in order to categorize them. For example, some students grouped integrals together in a group called "trig," others named a similar group "trig substitution," and still others "sines and cosines." For the most part, their intention was clear, but sometimes this resulted in ambiguous descriptions and a blurring of categories.

We note that the use of edit distance gives focus to another open question in learning. The standard definition of edit distance treats costs for making bonds and for breaking bonds equally. But it is unclear whether this is an accurate reflection of what is happening internally. It is possible that card sort tasks like this, in combination with microgenetic studies of the time involved could provide data to measure how these costs are related.

Finally, we note that while helpful for exploring student category development, the links-gaps graphs have some unusual features that complicate interpretation. Although one is tempted to apply traditional measures of distance to study changes in student learning, the space does not admit a simple distance measure. There are, in fact, a number of points in the space that cannot be reached. These holes in the space are not uniformly distributed throughout the space.

Conclusion

This study demonstrated two major results. First, we have shown that an alternative ordering of the traditional learning of techniques of integration, one focused on developing a larger conceptual structure of the types of integrals before learning specific procedural skills, can be effective. Second, the close examination of student categorization of integrals here gives quite a bit of insight into student thinking and learning about integration, and possibly about their developing understanding of other topics.

The approach to the teaching of techniques of integration here started by exploring how integrals are different from each other. Feedback allowed students to adjust their categories of integrals, incorporating information about the algebraic structure of the integrand, the result of the integration, and the relationship between integration and differentiation. After this experience, students studied each of the major types of integration in detail. While they performed adequately on the exam most proximal to the learning experiences, their work on the final exam on related material was quite a bit stronger. This is likely due to the time students had to reorganize their thinking and to incorporate all of their thinking into a more coherent schema. The key to this, we believe, is the way student learning was scaffolded, so that the students' minds could continue to reorganize the ideas even when not directly engaging with similar content.

In the traditional ordering of these materials, students are also encouraged to build a heuristic for integration, but this is done after they have already learned about the separate techniques. It remains to be seen whether students experiencing this ordering of content also have the ability to reorganize their thinking, even when not focusing on this content. We should also point out that most texts on differential equations begin with a chapter on categorizing differential equations before learning methods for solving them, following the learning process demonstrated here, although it is considerably easier to categorize differential equations by solution method without implementation.

With regard to the second aspect of this study, we demonstrate learning is not monotonic. Students often move further from "correct thinking" in the process of learning. Thus, being wrong is not the opposite of being right. Only by closer examination of the ways in which they have made mistakes can we truly understand student learning and help students move toward deeper understanding. Moreover, students vary greatly in the trajectories they take through the learning space on their way toward understanding, suggesting that variability is much greater than typically suspected. Students also require sufficient time for their schema to solidify after appropriate initial experiences. These implications may require the mathematics education community to rethink the efficacy of compressed-schedule summer courses in calculus, for example.

References

Abdul-Rahman, S. (2005). Learning with examples and students' understanding of integration. http://math.unipa.it/~GRIM/21_project/21_malasya_AbdulRahman24-28_05.pdf

Adey, P. (1999). *The science of thinking, and science for thinking: A description of Cognitive Acceleration through Science Education (CASE)*. Geneva, Switzerland: International Bureau of Education.

Alibali, M.W., Phillips, K.M.O., & Fischer, A.D. (2009). Learning new problem-solving strategies leads to changes in problem representation. *Cognitive Development, 24*, 89-101.

Anderson, M. (2010). *Wolfram Alpha for inquiry-based learning in calculus.* Accessed June 4, 2013 from http://busynessgirl.com/wolfram-alpha-for-inquiry-based-learning-in-calculus/

Barsalou, L. W. (1983). Ad hoc categories. *Memory and Cognition, 11*(3), 211-227.

Bennett, A., Moore, T., & Nguyen, X. (2011). *A longitudinal study on students' development & transfer of the concept of integration*. Paper presented at the Conference & Exposition, American Society for Engineering Education, Vancouver, British Columbia, Canada http://perg.phys.ksu.edu/reese/papers/2011/Bennett_ASEE2011_Paper-FINAL-SUBMITTED.pdf

Buchberger, B. (1990). Should students learn integration rules? *ACM SIGSAM Bulletin, 24*(1), 10-17. Retrieved from http://dl.acm.org/citation.cfm?id=1095228

Byrnes, J. P., & Wasik, B. A. (1991). Role of conceptual knowledge in mathematical procedural learning. *Developmental Psychology, 27*(5), 777-786.

Campbell, M. (2011). *Modeling the co-development of strategic and conceptual knowledge in mathematical problem solving,* unpublished dissertation.

Deibel, K., Anderson, R., & Anderson, R. (2005). Using edit distance to analyze card sorts. *Expert Systems, 22*(3), 129-138.

diSessa, A. (2001). *Changing minds: Computers, learning and literacy*. Cambridge, MA: Bradford Books.

Edwards, B. S., & Ward, M. B. (2004). Surprises from mathematics education research: student (mis)use of mathematical definitions. American *Mathematical Monthly, 111*(5), 411-424.

Fincher, S., & Tenenberg, J. (2005). Making sense of card sorting data. *Expert Systems*, *22*(3), 89-93.

Gordon, S. (1993). Calculus must evolve. *PRIMUS, 3*(1), 11-18. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/10511979308965683

Grandsard, F. (1997). On teaching strategies in mathematics. *The Mathematics Educator*, *2*(2), 166-180.

Gradshteyn, I. S., & Ryzhik, I. M. (1994). *Table of integrals, series, and products*. Academic Press: New York.

Green, K., & Ricca, B. (2013). The evolution of student ideas. *Proceedings for the 40th Annual Meeting of the Research Council on Mathematics Learning*, 189-196.

Green, K. H., & Ricca, B. (2014) Graph theoretic methods for the analysis of data in developing systems. *Quality and Quantity,* Retrieved from http://dx.doi.org/10.1007/s11135-014-0089-5.

Grundmeier, T., Hansen, J., & Sousa, E. (2006). An exploration of definition and procedural fluency in integral calculus. *PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies*, *16*(2): 178-191.

Harel, G. (2008). What is mathematics? A pedagogical answer to a philosophical question. In B. Gold, & R. Simons, (Eds.) *Proof and other dilemmas: Mathematics and philosophy* (pp. 265-290). MAA. Downloaded June 4, 2013 from http://math.ucsd.edu/~harel/Downloadable/What%20Is%20Mathematics.pdf

Harper, M., Jentsch, F., Berry, D., Lau, H. C., Bowers, C., & Salas, E. (2003). TPL-KATS-card sort: A tool for assessing structural knowledge. *Behavior Research Methods, Instruments, and Computers*, *35*(4), 577-584.

Heid, M. K. (1988). Resequencing skills and concepts in applied calculus using the computer as a tool. *Journal for Research in Mathematics Education*, *19*(1), 3-25.

Lauten, A. D., Graham, K., & Ferrini-Mundy, J. (1994). Student understanding of basic calculus concepts: Interaction with the graphics calculator. *Journal of Mathematical Behavior*, *13*, 225-237.

Mackie, D. (2002). Using computer algebra to encourage a deep learning approach to calculus. In *Proceedings of the 2nd International Conference on the Teaching of Mathematics (at the undergraduate level)*.

Available online http://server.math.uoc.gr/ ictm2/Proceedings/pap415.pdf, accessed Dec 17, 2012. http://server.math.uoc.gr/~ictm2/Proceedings/pap415.pdf

Mahir, N. (2009). Conceptual and procedural performance of undergraduate students in integration. *International Journal of Mathematical Education in Science and Technology*, *40*(2), 201-211.

Mallet, D. G. (2011). An example of cognitive obstacles in advanced integration: the case of scalar line integrals. *International Journal of Mathematical Education in Science and Technology*. Retrieved from http://eprints.qut.edu.au/50313/1/mallet_cogobs.pdf

McNeil, N. M., Rittle-Johnson, B., Hattikudur, S., & Petersen, L.A. (2010). Continuity in representation between children ad adults: Arithmetic knowledge hinders undergraduates' algebraic problem solving. *Journal of Cognition and Development, 11*(4), 437-457.

Meel, D. (1998). Honors students' calculus understandings: comparing calculus&mathematica and traditional calculus students. *CBMS Issues in Mathematics Education*, *7*, 163-215. Retrieved from http://personal.bgsu.edu/~meel/Research/meel(RCME1998).pdf

Moshtagh, N. (2005). Minimum volume enclosing ellipsoid. *Convex Optimization*. Accessed online June 5, 2013 from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.116.7691&rep=rep1&type=pdf

Murphy, L. (1999). Computer algebra systems in calculus reform. http://mste.illinois.edu/users/Murphy/Papers/CalcReformPaper.html

Oates, G. (2009). Relative values of curriculum topics in undergraduate mathematics in an integrated technology environment. In R. Hunter, B. Bicknell, & T. Burgess (Eds.), *Crossing divides: Proceedings of the 32nd annual conference of the Mathematics Education Research Group of Australasia (Vol. 2)*. Palmerston: North, NZ: MERGA. Retrieved from http://www.merga.net.au/documents/Oates_RP09.pdf

Orton, A., (1983). Students' understanding of integration. *Educational Studies in Mathematics*, *14*(1), 1-18.

Palmiter, J. (1991). Effects of computer algebra systems on concepts and skill acquisition in calculus. *Journal for Research in Mathematics Education*, *22*(2), 151-156.

Pettersson, K., & Scheja, M. (2008). Algorithmic contexts and learning potentiality: a case study of students' understanding of calculus. *International Journal of Mathematical Education in Science and Technology*, *39*(6), 767-784.

Real, R. (1999). Tables of significant values of Jaccard's index of similarity. *Micellania Zoologica, 22*(1), 29-40.

Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? Journal of *Educational Psychology, 91*(1), 175-189.

Rittle-Johnson, B., & Koedinger, K. R. (2005). Designing knowledge scaffolds to support mathematical problem solving. *Cognition and Instruction, 23*(3), 313-349.

Rosken, B., & Rolka, K. (2007). Integrating intuition: The role of concept image and concept definition for students' learning of integral calculus. In *TMME Monograph 3*: *(pp. 181-204)*. Retrieved from http://www.math.umt.edu/tmme/Monograph3/RoeskenRolka_Monograph3_pp.181_204.pdf

Schoenfeld, A. (1978). Presenting a strategy for indefinite integration. *American Mathematical Monthly*, *85*(8), 673-678.

Schroeder, L. (2011). Investigating a metacognitive strategy for solving indefinite integration problems in calculus: An fMAR study. Retrieved from http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED534621&ERICExtSearch_SearchType_0=no&accno=ED534621

Slagle, J. R. (1963). A heuristic program that solves symbolic integration problems in freshman calculus. In *Computers and Thought*, 191-203.

Sofronas, K., DeFranco, T., Vinsonhaler, C., Gorgievski, N., Schroeder, L., & Hamelin, C., (2011). What does it mean for a student to understand the first-year calculus? Perspectives of 24 experts. *Journal of Mathematical Behavior, 30*, 131-148.

Stewart, J. (2008). Single variable calculus: Early transcendentals, (6th Ed). Brooks/Cole Cengage Learning: Belmont, CA.

Tall, D. O. (1993). Students difficulties in calculus. *Proceeding of Working Group 3 on Students' Difficulties in Calculus*. ICME-7, Québec, Canada, 13-28.

Thelen, E., & Smith, L. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

Vygotsky, L. (1986). *Thought and language*. Cambridge, MA: MIT Press.