Building Models of Students' Use of Sigma Notation

[1]Kristen Vroom, [1]Sean Larsen, & [2]Stephen Strand

[1]Portland State University, [2]California State University, Chico

Summation notation is a widely-used standard that can represent all kinds of sums. Despite its utility, the literature on this topic points to the notation being difficult for students. Our research project gives insight into how students think about summation notation and why it is so challenging. This report builds off of the first phase in our project, which proved the existence of students' uncertainties with elements of the notation. Survey data from 285 undergraduates suggested that uncertainties are common amongst students. We also found that the act of encoding a sum in sigma notation is more cognitively demanding than interpreting a summation notation expression. In this paper we present models of students' ways of thinking about summation notation.

Keywords: Summation notation, calculus, cognitive models

Sigma Summation notation is a widely-used standard for expressing all kinds of sums, from infinite series to probabilities to approximate area under a curve. However, there is little in the research literature regarding this notation. There is consensus that the notation presents difficulties for students. For example, we were able to find papers discussing issues that contribute to student struggles in statistics (Ramsey, 1999) and engineering (Armstrong & Croft, 1999) that mention summation notation. Little and Jones (2010) provide empirical evidence that the notation is challenging for students and they observed that, "The use of algebraic language such as sigma notation and iteration formulae added substantially to the difficulty of questions" (p. 141). However, Little and Jones' study provides no information regarding why students might struggle with this notation or how they might come to make better sense of it.

Summation notation appears in some research articles on infinite series (e.g., Martínez-Planell, Gonzalez, DiCristina, & Acevedo, 2012) and integration (e.g., Sealey, 2014) but is not the focus of these studies. Martinez-Planell et al., focus on the distinction between seeing an infinite series as an infinite sum and seeing it as the limit of a sequence of partial sums. They do not investigate possible challenges involving the sigma notation used to represent partial sums and infinite series. Sealey presents a framework for understanding Riemann sums that includes a "summation layer" but acknowledges that her study, "did not require students to use the notation of $\sum_{i=1}^{n} f(x_i)\Delta x$ to represent the Riemann sum, nor did it explore students' understanding of this notation" (p. 243). Additionally, Brijlall and Bansilal (2010) proposed a genetic decomposition of the Riemann sum that explicitly attended to summation notation. Specifically, they included two abilities related to summation notation in their model proposing that students who understand Riemann sums 1) "can represent finite or infinite sums as expanded sums, when given its compact form using sigma notation" and 2) "can represent finite or infinite sums in compact form using sigma notation when given its expanded representation." (p. 133-134). While this study provides no information about whether (and if so, why) students struggle with these two activities, this pair of activities provides a structured way to think about what it may mean for students to be able to make sense of and work fluently with this notation.

What little information exists suggests that students struggle with summation notation and that this likely contributes to difficulties with concepts that rely on this notation such as probability, infinite series, and integration. However, the research literature contains very little insight into *how* students think about this notation and *why* it is challenging. Our research project endeavored to investigate how students think about summation notation and was completed in two phases.

In the first phase, summation notation was not the focus. We conducted an exploratory design experiment with two students, Betty and Kathy, where we actively investigated post-calculus students' understanding of integral. During the experiment, summation notation came up and the students displayed some interesting thinking about how the notation functioned. When we investigated their thinking we found that Kathy and Betty were aware that sigma notation provides a shorthand way to write a sum. They seemed keenly aware of the kinds of information needed to expand a sum (or that should be captured in the notation when encoding a sum). However, they were unsure exactly how this information was supposed to be recorded in the notation. For example, they were uncertain whether the value on top of the sigma was supposed to represent the terminal input value or the number of terms (Strand, Zazkis, & Redmond, 2012).

After working with Betty and Kathy we wondered if their uncertainties are common amongst other undergraduate students. For this reason we designed a second phase specifically targeting student thinking about summation notation. The results from the second phase of our project will be the focus of this paper. We aim to address the following research questions: (1) If students do struggle with sigma notation, what kinds of difficulties do they have and how might such difficulties be explained? (2) Is there a difference in difficulty between expanding (interpreting) sums expressed in summation notation and compressing (encoding) expanded summations using summation notation?

Methods

Survey Instrument

Our survey instrument consisted of three tasks. The first task involved encoding a Riemann sum, which required the student to express the input variable of a function as a function of the index variable. The second task was a basic encoding task in that the index variable could serve as the input variable. The third task was an interpretation task. We conjectured the first task was the most difficult and the third task to be the least difficult. This study focuses on student responses of the second and third tasks. Here we will describe the two tasks (in reverse order) and the rationale for their design.

The Interpreting Task (Task 3) instructed students to write out the expanded ("longhand") sum represented by a given sigma notation expression (Figure 1). We expected a student adhering to the standard convention of summation notation to write, "$(2+1)^2 + (3+1)^2 + (4+1)^2 + (5+1)^2 + (6+1)^2 + (7+1)^2$." Specifically, the "2" below the sigma represents the starting value of the index, the "7" above the sigma represents the terminal value of the index, and each step in the index is incremented by 1. We chose to start the index at two so that we would be able to tell if the seven above the sigma sign was interpreted as the number of terms or the terminal value of the index. Betty and Kathy vacillated between these two interpretations of the index and so we wished to know how common this particular difficulty was. In general we wished to see what kinds of errors students might make in interpreting the different elements of summation notation (the index, the sigma, the summand, the numbers above and below the sigma).

Task 3: Write out the given summation longhand:

$$\sum_{k=2}^{7}(k+1)^2$$

Figure 1. The Interpreting Task.

The Encoding Task (Task 2) asked students to encode the sum of the first ten odd integers using summation notation (Figure 2). There are many possible expressions that would follow the standard convention of summation notation but an example would be “$\sum_{k=1}^{10} 2k-1$ ” or possibly “$\sum_{k=0}^{9} 2k+1$”. With this task we were interested in what challenges encoding with summation notation would present to the students. We were also interested in exploring the relative difficulty of encoding and interpretation tasks.

Task 2: Using $\sum$-notation, write an expression for the sum of the first ten odd integers.

Figure 2. The Encoding Task.

The Interpreting and Encoding Tasks were exactly the same as tasks given to Betty and Kathy during the design experiment. We chose to give these tasks to Betty and Kathy so that we could investigate what kinds of errors students might make with summation notation in less complex contexts than the first task. The consistency across the two phases of the projects allowed us to leverage Betty and Kathy’s reasoning when we analyzed the survey data.

Participants

Two hundred eighty five undergraduate students participated in the second phase of our study. These students were enrolled in a course in the calculus sequence, differential equations, linear algebra, or a course in an undergraduate advanced calculus sequence and received extra credit for their participation. We invited students enrolled in this set of courses in hopes of receiving responses from students with various summation notation experiences. In total, 567 students were enrolled in at least one of 15 sections of the 8 courses. 50.3 percent (285/567) of the students completed at least part of the survey, 42.3 percent (242/567) of the students completed both Task 2 and Task 3.

Analysis

There are four stages of analysis for the Interpreting and Encoding Tasks. During the first stage we went through each survey and counted how many had errors of any kind; these were sorted by task. In the first pass we looked to see if each solution was perfect; if it was then the response was marked ‘correct’ and if not it was marked as ‘incorrect’. We did not attempt to analyze or describe the errors at that stage.

In the second stage of analysis we recorded error types made on the Interpreting Task. To do so we first read each survey’s responses and discussed common errors. We came up with 6 categories of error types, including: gave a sum with seven summands, substituted only even values for k, not enough summands, used sigma with the expanded sum, used only one value as the input, and other.

In the third stage we turned to the Encoding Task. While coding error types for the Interpreting Task involved low inference, we felt this would not be the case for the Encoding Task. For instance, one student produced the following expression: "$\sum_{i=0}^{10} i+1$" (Figure 4). We could conjecture that the student incorrectly substituted index values starting at $i=0$ and ending at $i=10$; however, this would rely heavily on our interpretation. Later, we will argue that this is probably not how this student would explain their encoded sum. In order to match the level of inference used to code the Interpreting Task, we recorded each piece of the given summation expression. For instance, we recorded what (if anything) was written below and above the sigma. We also recorded the type of expression within the sigma into the following categories: outputs odd values when incrementing the inputs by 1 (e.g., $2x-1$), outputs odd values when not incrementing the inputs 1 (e.g., $x+1$), and seemingly unproductive expression (e.g., $\frac{(-1)^x}{x}$). In this stage we also sorted responses that we tagged as 'correct' into three categories: correct and summed $2i+1$ (or equivalent) from $i=0$ to $i=9$, correct and summed $2i-1$ (or equivalent) from $i=1$ to $i=10$, or correct with other encoding.

The last stage of analysis involved comparing each student's Interpreting Task response to the Encoding Task response. This process included separating surveys based on error type and then looking at the frequency of the Encoding Task tags. For example, we found that 16 of the 27 students that gave an expression with seven summands in the Interpreting Task also wrote a "10" above the sigma in the Encoding Task.

Sample Results

A Model for Interpreting Summation Notation

We found that most students were able to answer the Interpreting Task correctly; 73.2 percent (186/254) of the students that attempted the task correctly answered the question. We will first attempt to explain the thinking of the 69 students that incorrectly answered the Interpreting Task. We found the most common error type (27 responses) involved summing seven terms. There were 4 different subcategories within this error type, which are listed with their frequency in Table 1.

Table 1. Categories for summing seven terms.

Category	Frequency (n=27)
Summed $(k+1)^2$ from $k=2$ to $k=8$	16
Substituted 2 seven times	6
Did not increment by 1	3
Did not substitute values for k	2

Summing seven terms suggests that these students took the number atop the sigma as the number of summands. Kathy and Betty also considered this during the first phase of the project. While this is not the standard convention (i.e., the value above the sigma represents the terminal index value), it *is* a viable convention and it is analogous to a "count loop" in computer science.

However, the non-standard usage is still problematic because a user who respected the standard usage would be unable to recreate the expanded sum the students meant to encode.

We can test whether or not the 27 students truly take the number above the sigma to be the number of summands by considering their responses to the Encoding Task. The 27 students should write “10” above the sigma when they are prompted to encode the sum of the first *ten* odd integers if they are using the nonstandard convention.

The following section will investigate how students who ‘summed $(k+1)^2$ from $k=2$ to $k=8$’ responded to the Encoding Task.

Subcategory 1: Summed $(k+1)^2$ from $k=2$ to $k=8$. Figure 3 gives a typical response of the 16 students under the “summed $(k+1)^2$ from $k=2$ to $k=8$” subcategory. Responses in this subcategory include 7 summands, each of which correspond to an output of the function $f(k)=(k+1)^2$ with $\{2,3,\dots,8\}$ as its domain.

$$(2+1)^2 + (3+1)^2 + (4+1)^2 + (5+1)^2 + (6+1)^2 + (7+1)^2 + (8+1)^2$$

Figure 3. Typical Interpreting Task response with ‘summed $(k+1)^2$ from $k=2$ to $k=8$’ tag.

Now consider an Encoding Task response from one of the students that made a subcategory 1 error in the Interpreting Task (Figure 4). Notice, this student did indeed place a “10” above the sigma. In total, 10 of the 16 responses did so as well.

$$\sum_{i=0}^{10} i+1$$

Figure 4. Student response to Encoding Task.

Two of the 6 students that did not place a “10” above the sigma left the Encoding Task (seemingly) incomplete; both responses did not include an expression to sum. Additionally, it appears as though another student misread the Encoding Task to say ‘write an expression for the sum of odd numbers less than ten.’ (See figure 5.) This is evidenced by the list “{1, 3, 5, 7, 9}”. Under this assumption, the number above the sigma, “6”, corresponds to the number of odd values less than 10.

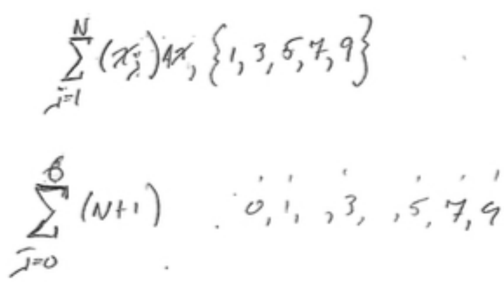

Figure 5. Possible misinterpretation of Encoding Task prompt.

It is also interesting to note that it is possible for the students to assign the number above the sigma to mean the total number of summands *and* give a ‘correct’ answer to the encoding task. In this situation, a student could interpret “$\sum_{k=1}^{10} 2k-1$” to mean: substitute natural numbers for

k until there are 10 summands. We found 3 of the 10 student responses that placed a "10" above the sigma was tagged as 'correct'.

However, this convention is problematic whenever the starting variable is not equal to 1. In this situation, the number of summands does not equal the number of values to substitute using the standard convention. There were 40 Encoding Task responses that were tagged 'correct' because they gave the expression "$\sum_{k=0}^{9} 2k + 1$". This response suggests that these students do *not* take the number above the sigma to mean the number of summands since the prompt asked students to sum the first *ten* odd numbers. We found that only 1 of the 40 students that answered "$\sum_{k=0}^{9} 2k + 1$ " to the Encoding Task also wrote seven summands for the Interpreting Task.

Task Hierarchy

While most students were able to correctly answer the Interpreting Task, only 38.4 percent (96/250) of the students that attempted the Encoding Task wrote a correct response. We conjectured that encoding is more cognitively demanding than interpreting a summation-notation expression. This is because if a student were able to encode a sum, then the student would be familiar with the structure and workings of the elements of the summation notation; that is, the student would have the ability to accurately interpret a given summation-notation expression. Additionally, in order to correctly verify their encoding, a student would necessarily be able to interpret their own summation notation accurately.

With respect to the Tasks we expected this to play out in a contrapositive manner. That is, we expected that students who could not successfully complete the Interpreting Task would not be able to successfully complete the Encoding Task. Table 2 shows the number of students that correctly answered both tasks correctly, neither of the tasks correct, Interpreting Task correctly but Encoding Task incorrectly, and Encoding Task correctly but Interpreting Task incorrectly.

Table 2. Combination of correct responses to Interpreting and Encoding Task.

Tasks Correct	Frequency (n=242)
Both Interpreting and Encoding Tasks	86
Neither Interpreting nor Encoding Tasks	54
Only Interpreting Task	93
Only Encoding Task	9

Notice, of the 64 participants that were not able to answer the Interpreting task correctly, 9 participants were able to answer the Encoding task correctly. This seemed to be a high percentage until we looked further into how the 9 students responded. In particular, 4 of these students attempted to answer the Interpreting Task by expanding $(k+1)^2$. One of these students then made an algebraic error (unrelated to the notation). The 3 other students distributed the sum across the three terms, evaluated the first two summations correctly, and then incorrectly reduced "$\sum_{k=2}^{7} 1$ " to "$1 \cdot 7$" (Figure 6). It seems as though evaluating the sum of a constant function over an index set might be more difficult for students since the students were able to correctly evaluate the first two terms after they distributed the sum.

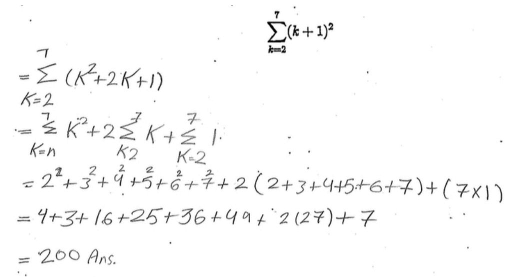

Figure 6. Student 'incorrect' response to Interpreting Task.

An additional 2 students answered the interpreting task incorrectly by summing seven terms. These students then answered the Encoding task by writing "$\sum_{k=1}^{10} 2k - 1$". This is a situation that we described before in the previous section. That is, it is likely that these students interpret the number above the sigma to mean the number of summands, not the terminal index value. When students assign the initial value of the index to be 1 (and increment by 1) both ways of thinking will produce the same expression.

We also hypothesized that the ability to interpret a summation notation expression would not be sufficient for successfully encoding with the notation. This is because encoding sums with summation notation requires an understanding of functions and their domains above and beyond what is required to accurately interpret a given summation-notation expression. Specifically one must *construct* a function; this entails coordinating the construction of a rule with the construction of an appropriate domain (i.e. indexing set, in this case). For example in the Encoding Task, a student must first construct a function that will output odd values when the student inputs natural numbers (e.g., $f(x) = 2x - 1$). Then the student must restrict the natural numbers in such a way that the indexing set will output the first ten odd numbers (e.g., $\{1, 2, 3, \dots, 10\}$). In contrast, to interpret a student does not have to coordinate the construction of a rule with the construction of an appropriate domain. Instead, the student only coordinates the inputs from the given index set with the outputs of a given function. For this reason we expected a number of students to successfully complete the Interpreting Task who would not successfully complete the Encoding Task. Table 1 shows that this was indeed the case: 93 participants were able to successfully complete the Interpreting Task but were not able to complete the Encoding Task.

Conclusion

We found that undergraduates struggle with summation notation. Like Betty and Kathy, many students were unsure about the structure and workings of the elements of summation notation. In particular, our data suggest that some students are unaware of (or at least do not follow) the standard summation notation conventions. However, often students do follow a non-standard convention (e.g., assign the number above the sigma to mean the number of summands, increment by a value other than 1, etc.). The conventions we witnessed are viable options; however, they are problematic because they are not standard. In particular, issues may arise when communicating with others that adhere to the standard convention.

We also found that encoding is more cognitively demanding than interpreting a summation-notation expression. We presented evidence supporting that being able to expand sums expressed in summation notation is necessary to being able to encode an expanded sum using summation notation. However, the ability to interpret an expression in sigma notation is not enough to encode a sum in sigma notation.

References

Armstrong, P. K., & Croft, A. C. (1999). Identifying the learning needs in mathematics of entrants to undergraduate engineering programmes in an English university. *European Journal of Engineering Education*, *24*(1), 59-71.

Brijlall, D., & Bansilal, S. (2010). A genetic decomposition of the Riemann Sum by student teachers. In *Proceedings of the eighteenth annual meeting of the South African Association for research in mathematics, science and technology education* (pp. 131-138).

Little, C., & Jones, K. (2010). The effect of using real world contexts in post-16 mathematics questions. In *Proceedings of the British Congress for Mathematics Education April 10* (pp. 137-144).

Martínez-Planell, R., Gonzalez, A. C., DiCristina, G., & Acevedo, V. (2012). Students' conception of infinite series. *Educational Studies in Mathematics*,*81*(2), 235-249.

Ramsey, J. B. (1999). *Proccedings of the 52th Session of the ISI. Helsinki*, 10-18.

Sealey, V. (2014). A framework for characterizing student understanding of Riemann sums and definite integrals. *The Journal of Mathematical Behavior*,*33*, 230-245.

Strand, S., Zazkis, D., & Redmond, S. (2012). Summing up student's understanding of sigma notation. *Presentation at the 15th Research in Undergraduate Mathematics Education Conference.* Portland, OR.