Classifying combinations:
Do students distinguish between different types of combination problems?

Elise Lockwood	Nicholas H. Wasserman	William McGuffey
Oregon State University	Teachers College, Columbia University	New York, NY

In this paper we report on a survey designed to test whether or not students differentiated between two different types of problems involving combinations – problems in which combinations are used to count unordered sets of distinct objects (a natural, common way to use combinations), and problems in which combinations are used to count ordered sequences of two (or more) indistinguishable objects (a less obvious application of combinations). We hypothesized that novice students may recognize combinations as appropriate for the first type but not for the second type, and our results support this hypothesis. We briefly discuss the mathematics, share the results, and offer implications and directions for future research.

Key words: Combinatorics, Discrete mathematics, Counting

Discrete mathematics, with its relevance to modern day applications, is an increasingly important part of students' mathematical education, and prominent organizations have called for increased teaching of discrete mathematics topics in K-16 mathematics education (e.g., NCTM, 2000). Combinatorics, and the solving of counting problems, is one component of discrete mathematics that fosters deep mathematical thinking but that is the source of much difficulty for students at a variety of levels (e.g., Batanero, Navarro-Pelayo, & Godino, 1997; Eizenberg & Zaslavsky, 2004). The fact that counting problems can be easy to state but difficult to solve indicates that there is a need for more research about students' thinking about combinatorics.

One fundamental building block for understanding and solving combinatorial problems are combinations (i.e., $C(n,k)$, also called binomial coefficients due to their role in the binomial theorem). Combinations are prominent in much of the counting and combinatorial activity with which students engage, and yet little has been explicitly studied with regard to student reasoning about combinations. This study contributes to our understanding of students' reasoning about combinations, and in particular to study beginning students' inclination to differentiate between typical combinatorics problems. This study addresses the following research question: *Do early undergraduate students recognize two different types of combination problems as involving binomial coefficients, and do they use binomial coefficients to solve both types of problems?*

Literature and Theoretical Perspective

Combinations in Mathematics Education Literature. As we have noted, there is much documented evidence for the fact that students struggle with solving counting problems correctly. Some reasons for such difficulty are that counting problems are difficult to verify (Eizenberg & Zaslavsky, 2004) and that it can be difficult to effectively encode outcomes[1] in terms of objects one knows how to count (e.g., Lockwood, Swinyard, & Caughman, 2015b). We seek to address potential difficulties by focusing on better understanding students' application and use of combinations – which are fundamental in enumeration. Piaget & Inhelder (1957) studied students' mental processes as they solved arrangement and selection problems, and they

[1] We take *encoding outcomes* to be the combinatorial activity of articulating the nature of what is being counted by associating each outcome with a basic mathematical entity (such as a set or sequence).

took a special interest in determining whether permutations, arrangements, or combinations would be most difficult for students. Dubois (1984), Fischbein and Gazit (1988), and Batanero, et al. (1997) have also investigated the effects of both implicit combinatorial models and particular combinatorial operations on students' counting, again considering differences in reasoning about particular problem types such as permutations and combinations. We extend existing work that focuses on students' mental processes of foundational combinatorial ideas, seeking specifically to explore the extent to which undergraduate students distinguish between two types of combination problems (which we develop in the following section). Our work also builds on a recent study by Lockwood, Swinyard, & Caughman (2015a) in which two undergraduate students reinvented basic counting formulas, including the formula for combinations. Based on the students' work on combination problems, Lockwood, et al., (2015b) suggested the importance of being able to correctly encode outcomes combinatorially (by which they mean the act of articulating the nature of what is being counted by associating each outcome with a mathematical entity such as a set or a sequence).

Mathematical Discussion. A combination is a set of distinct objects (as opposed to a permutation, which is an arrangement of distinct objects). Combinations can also be described as the solution to counting problems that count "distinguishable objects" (i.e., without repetition), where "order *does not* matter." The total number of combinations of size k from a set of n distinct objects is denoted C(n,k) and is verbalized as "n choose k."[2] As an example, combinations can be used if you want to select from eight (distinguishable) books three books to take on a trip with you (order does not matter) – the solution is C(8,3), or 56 possible combinations. By contrast, other combinatorial problems and solution methods, such as permutations, are organized in relation to some different possible constraints – see Table 1.

Table 1: Selecting *k* objects from *n* distinct objects

	Ordered	**Unordered**
Distinguishable Objects (without repetition)	Permutations $\frac{n!}{(n-k)!} = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot (n-k+1)$	Combinations $\binom{n}{k} = \frac{n!}{k!(n-k)!}$
Indistinguishable Objects (with repetition)	Sequences $n^k = \underbrace{n \cdot n \cdot n \cdot \ldots \cdot n}_{k}$	Multicombinations $\left(\binom{n}{k}\right) = \binom{k+n-1}{n-1}$

In this paper we refer to *combination problems* as problems that can be solved using binomial coefficients, in the sense that parts of their outcomes can be appropriately encoded as sets of distinct objects (i.e., Lockwood, et al., 2015b). Sometimes this encoding is fairly straightforward, as the outcomes are very apparently sets of distinct objects. For instance, in the Basketball problem (stated in Table 2), the athletes could be numbered 1 through 12 (because they are different people), and the outcomes are fairly naturally modeled as 7-element sets taken from the set of 12 distinct athletes. Any such set is in direct correspondence with a desired outcome; there are C(12,7) of these sets. We call such problems Type I. In other situations, a problem may still appropriately be solved using a combination, but recognizing how to encode the outcomes as sets of distinct objects is less clear. For example, consider the Coin Flips problem (stated in Table 2). Here, the problem also can be solved using combinations – that is, if the outcomes are encoded appropriately. The most natural way to model an outcome is as an ordered sequence of Hs and Ts; however, we can encode a desirable outcome as a set of distinct

[2] The derivation of the formula for $C(n,k)$ as $n!/(n-k)!k!$ is not pertinent to the study; Tucker (2002) provides a useful explanation.

positions in which the Hs are placed. Given the five possible positions (i.e., the set: {1, 2, 3, 4, 5}), the outcome HHTHT would be encoded as the set {1, 2, 4}. This sufficiently establishes a bijection between outcomes and sets because every outcome has a unique placement for the Hs (the Ts must go in the remaining positions). In this way, the answer to the counting problem is simply the number of 3-element subsets from 5 distinct objects (i.e., positions 1-5), which is C(5,3). We call these Type II problems (See Table 2).

Table 2: Characterizing two different "types" of fairly standard combination problems

	Description	Example problem	Natural Model for Outcomes
Type I	*An unordered selection of distinguishable objects*	*Basketball Problem.* There are 12 athletes who try out for the basketball team – which can take exactly 7 players. How many different basketball team rosters could there be?	{(1,2,3,4,5,6,7), (1,3,5,7,9,11,12), …}
Type II	*An ordered sequence of two (or more) indistinguishable objects*	*Coin Flips Problem.* Fred flipped a coin 5 times, recording the result (Head or Tail) each time. In how many different ways could Fred get a sequence of 5 flips with exactly 3 Heads?	{(HHTTH), (HTHHT), (TTHHH), …}

In light of various ways of encoding outcomes that facilitates the use of combinations, we point out that it may seem that combinations are actually being used to solve two very different kinds of problems. The outcomes in the Basketball problem are clearly unordered sets of distinct objects, but the outcomes in the Coin Flips problem are actually *ordered* sequences (not unordered) of two kinds of *indistinguishable* (not distinct) objects (Hs and Ts). Combinations are applicable in both situations, but we argue that there could be a difference for students in identifying both problems as counting combinations. Indeed, using combinations to solve Type II problems involves an additional step of properly encoding the outcomes with a corresponding set of distinct objects, and we thus posit that Type I problems would be more natural for novice students, more clearly representative of combination problems than Type II problems.

In spite of the widespread applicability of combinations, we posit that students may not recognize all fairly standard combination problems as involving combinations. This may be due in part to the fact that students tend not to reason carefully about outcomes (e.g., Lockwood, et al., 2015b), and because "distinguishable" and "unordered" are not always natural or clear descriptions of the situation or outcomes. We are thus motivated us to investigate whether or not students actually respond differently to the two different problem types.

Methodology

We designed two versions of a survey, and although the survey contained a number of elements, we focus in particular on features of the survey that serve to answer the research question stated above. Each survey consisted of 11 combinatorics problems, and each problem was designed with categories in mind that included problem type (I or II) and complexity (Simple, Multistep, or Dummy).[3] *Simple* combination problems refer to those that can be solved using a single binomial coefficient, in the sense that their outcomes can be appropriately encoded as sets of distinct objects; *multistep* combination problems would require multiple binomial coefficients in the solution. The authors coded the problems independently before finalizing the coding for each problem. Each version of the survey contained the same number of problems of

[3] We also coded the tasks according to other criteria that we do not report on in this paper, such as: sense of choosing (Active or Passive), and whether an object or process is to be counted (Structural or Operational).

each type and complexity, as well as two "dummy" problems to discourage students from assuming that every problem could be solved with a combination. Each problem was selected for one version of the survey with a companion problem in mind for the other version in order to compare responses with respect to the various coding categories.

We targeted Calculus students because they were believed to have been likely to have seen combinations at some point in their mathematical careers without having studied them in detail. In order to gain insight into students' familiarity with combination problems, each survey included a section on demographic information to identify previous mathematics courses taken in high school or in college, current mathematics courses being taken, and whether students recognized different representations of combinations and permutations. We have yet to incorporate the demographic data in our analysis, but this is a further avenue we plan to pursue.

In order to investigate the research question, we wanted to see whether or not students who would solve Type I problems using combinations would also solve Type II problems using combinations. That is, we wanted to see whether students would recognize a difference between Type I and Type II problems in terms of the applicability of combinations as a solution. In order to do this, we needed to focus on only those participants who had demonstrated some understanding of combinations, by using them in their solution to Type I problems. On each survey there were four *simple* Type I problems and three *simple* Type II problems. These seven problems were of particular interest. Two other problems, one *multistep* Type I and another *multistep* Type II, were also included, for which the answer was a product of combinations. We treated these two kinds of problems separately in the analysis.

The prompts for the combination problems asked for the students to use notation that suggests their approaches rather than numerical values,[4] but many students did not, on the whole, follow this prompt. Although there were 69 complete responses (we removed participants who did not finish the survey), 38 gave strictly numerical responses rather than expressions that indicated their solution method, which limited our ability to analyze their responses. Further, of the 31 remaining, only 12 correctly responded to at least half of the Type I problems using the appropriate combination notation. For the purpose of our preliminary report, we focus our analysis on these 12 participants, because, as noted, we sought to examine students who had used combinations to solve Type I problems. It is perhaps noteworthy that so few students followed the survey directions and also that so few correctly solve the Type I problems using combinations, but these are not points of discussion we are able to discuss in detail.

Findings

In this section we present two aspects of our data analysis that support a singular finding in regard to our research question.

Simple Type I and Type II problems. Out of the 12 people who correctly answered at least half of the Type I problems using the appropriate combination notation, 6 of these participants demonstrated very different responses between Type I and Type II problems. This is a considerable proportion (50%) of participants that displayed a dichotomous way of responding to these two problem types. For example, one participant gave correct responses of C(20,4), 9!/(7!2!), C(15,2), and C(250,6) for the Type I problems, and similarly correct but different

[4] Specifically, the prompt was: Read each problem and input your solution in the text box. **Please write a solution to the problem that indicates your approach.** If you're not sure, input your best guess. NOTE: Appropriate notation includes: 9+20, C(5,2), 5C2, 21*9*3, 5*5*5*5=5^4, 8!, 8!/5! = 8*7*6 = P(8,3), C(10,2)*3, Sum(i,i,1,10), 12!/(5!*7!), etc. *Only* if you individually count *all* of the outcomes should you input a numerical answer, such as 35.

responses of 8!/(5!3!), 55!/(20!35!), and 40!/(17!23!) for the Type II problems. In other words, even though this participant identified 3 of the 4 Type I problems as easily solvable by a combination, s/he did not recognize any of the Type II problems as a combination. Another participant gave correct responses of C(20,4), C(9,2), C(15,2), and C(250,6) for the Type I problems, but incorrectly gave permutation responses of P(13,2), P(55,2), and 40P(2,1) to the Type II problems. Again, this participant's responses indicate a difference between how s/he understood and solved these problems. Only two participants used combinations (correctly) to solve all of the Type I and Type II problems. Some other participants were mixed. In these cases, the Type I/II distinction does not seem to explain their responses as clearly and we hope further analysis of the data will lend additional insight into their responses.

Multistep Type I and Type II problems. As further support of the potentially differential responses between the two hypothesized problem types, we look at two multistep problems. Such problems may themselves be different given that participants have to view combinations as part of a process for arriving at the solution instead of the solution itself (like in simple combination problems). Regardless, for the multistep Type I problem, 10 of the 12 participants used some sort of combination in their solution. Although about half still arrived at an incorrect solution (for example, adding two combinations), the majority of these participants viewed combinations as vital to their solution approach. In contrast, for the multistep Type II problem, only 1 of the 12 participants used a combination in their solution at all (this participant, however, still arrived at an incorrect solution). The only participants with correct solutions – there were five of them – were of the form: 17!/(3!4!2!8!). Thus, we see further evidence in these multi-step combination problems to suggest that students seem to differentiate between these two problem types – viewing Type I problems as suitably involving combinations in the solutions whereas Type II problems require some other solution approach.

Conclusions and Implications

Despite the fact that all of these problems would be considered fairly standard combination problems, our findings suggest that the participants did not view the problems in this way. In particular, the 12 participants were mostly successful in using combinations to solve Type I problems, but often relied on other (at times incorrect) methods to solve Type II problems. These findings, while preliminary, suggest potential implications for the teaching and learning of combinations. They seem to indicate that students may not necessarily view the two problem types as the same, and perhaps with good reason: the descriptions of "unordered" and "distinct" do not seem to apply – at least in the most natural way to model the outcomes. Thus, students may need additional exposure to combinations and may benefit from explicit instruction about how Type II problems can be encoded in a way that is consistent with Type I problems. Generally, this point underscores a need for students to become more adept at combinatorial encoding (Lockwood, et al., 2015b). Encoding outcomes as sets is an inherent part of the field of combinatorics, but students may need particular help in making this connection explicit. Combinations are a powerful tool for enumeration problems, but without a robust understanding – including how and why Type II problems can be solved using them – students may possess a tool they do not really understand how to use. In addition, there are natural next steps and avenues for further research. We plan to investigate more questions and hypotheses with the data we have, such as analyzing effects of demographic data and investigating other relationships and potentially contributing factors in students' responses. Our findings also indicate that further investigating students' reasoning about encoding with combinations through in-depth interviews may give insight into the development of more robust understandings.

References

Batanero, C., Navarro-Pelayo, V., & Godino, J. (1997). Effect of the implicit combinatorial model on combinatorial reasoning in secondary school pupils. *Educational Studies in Mathematics, 32*, 181-199.

Dubois, J. G. (1984). Une systematique des configurations combinatoires simples. *Educational Studies in Mathematics, 15*(1), 37-57.

Eizenberg, M. M., & Zaslavsky, O. (2004). Students' verification strategies for combinatorial problems. *Mathematical Thinking and Learning, 6*(1), 15-36.

Fischbein, E., & Gazit, A. (1988). The combinatorial solving capacity in children and adolescents. *ZDM, 5*, 193-198.

Kavousian, S. (2008) *Enquiries into undergraduate students' understanding of combinatorial structures.* Unpublished doctoral dissertation, Simon Fraser University – Vancouver, BC.

Lockwood, E., Swinyard, C. A., & Caughman, J. S. (2015a). Patterns, sets of outcomes, and combinatorial justification: Two students' reinvention of counting formulas. *International Journal of Research in Undergraduate Mathematics Education, 1*(1), 27-62. Doi: 10.1007/s40753-015-0001-2.

Lockwood, E., Swinyard, C. A., & Caughman, J. S. (2015b). Modeling outcomes in combinatorial problem solving: The case of combinations. To appear in T. Fukawa-Connelly, N. Infante, K. Keene, and M. Zandieh (Eds.), *Proceedings for the Eighteenth Special Interest Group of the MAA on Research on Undergraduate Mathematics Education.* Pittsburgh, PA: West Virginia University.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: NCTM.

Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children*. New York: W. W. Norton & Company, Inc.

Tucker, A. (2002). *Applied Combinatorics* (4th ed.). New York: John Wiley & Sons.