Logical Reasoning in Probability: Further Investigation Into Students' Beliefs About the Relationship Between Covariance and Independence

Megan Ryals
University of Virginia

Brian Rickard
University of Arkansas

Morgan Sellers
Colorado Mesa University

In many undergraduate math courses, students are expected to interpret and use conditional statements without prior training in logical inference. In this paper we report on a portion of the third phase of a project exploring students' interpretations of a conditional statement in Probability relating covariance and independence. Specifically, we discuss the coding and analyzing of student written work in class and on assessments. We explain how this work was influenced by prior phases of the study. Quantitative results show no statistical difference between responses of students with and without prior training in logic. We do, however, see differences in students' responses overall as the semester progresses.

Keywords: Conditional Statements, Logical Reasoning, Probability, Covariance, Independence

Many theorems and definitions in undergraduate mathematics are stated as conditional statements. Consequentially, several researchers have begun to investigate how undergraduate mathematics students treat concepts from logic before they have taken a formal logic course (Case, 2015; Durst & Kaschner, 2020; Sellers et al., 2021). These studies have focused on students' meanings for conditional statements in calculus courses. In this paper, we report on an ongoing investigation into student interpretations of the following theorem from an undergraduate Probability course: "If X and Y are independent random variables, then Cov(X,Y)=0" (Casella & Berger, 2002, p. 171). Specifically, *we ask how students interpret variants of this theorem and how prior logic training impacts their interpretations.*

Literature Review

In this study, we focus on a conditional statement found in a theoretical probability course. Consequentially, we relied on literature related to both what is known about students' meanings for conditional statements as well as what is known about common hurdles that teachers might encounter when teaching an undergraduate probability course.

Students' Interpretations of Conditional Statements

A robust amount of previous literature suggests that students often interpret conditional statements as biconditionals (Hoyles & Küchemann, 2002; O'Brien et. al, 1971; Wason, 1968). Ryals et al. (2023, 2024) found that students in probability courses do often interpret the given theorem as a converse. Even if students have taken (or are currently taking) a logic course, students may not transfer principles of logic across different mathematical contexts; instead, their inferences depend on the context, particularly when it is a familiar one. (Dawkins & Cook, 2017; Durand-Guerrier, 2003; Sellers, 2020). In our study, we focus on whether students rely on rules of logic, their meanings for independence and covariance, or both in their interpretation of the theorem relating independence and covariance.

Students' Meanings for Probability Concepts

Some studies in undergraduate mathematics and statistics education investigate students' understandings of probabilistic ideas within statistics courses (Batanero, et al., 2004, 2016;

Pfannkuch & Constance, 2017). Kolmogorov's (1950) derived axiomatic theory of probability is a primary focus in undergraduate courses in probability. However, many of these studies relate to statistics education specifically and/or K-12 related domains, and little is known about how these undergraduate students think about many of the theorems like the one in our study. We do know from the mathematics and statistics education literature that many words in statistics and probability have "lexical ambiguities" (Barwell, 2005). These words may have different meanings in colloquial language than they do in mathematics and statistics classrooms (Sellers, 2020, 2024). Our study both draws from undergraduate mathematics education research in logic and what is known about lexical ambiguities to explore how undergraduate mathematics students might conceive of concepts in a theoretical probability course.

Theoretical Perspective

In this paper, we utilize a constructivist perspective (von Glasersfeld, 1995). We posit that each individual is constructing their own meanings for conditional statements as a whole, as well as constructing meanings for mathematical terms such as independence and covariance found within these definitions. We use the term meaning in the sense of Thompson et al. (2014), where a meaning is a scheme (Piaget, 1977). In light of this perspective, we focus on individual student responses related to independence and covariance first before categorizing commonalities in meanings across different students. In this phase of the study, we focus on our quantitative findings related to students' written work for the theorem. At this stage, we want to classify the characteristics of their argumentation (their justification for whether two variables are independent) in order that we might later use these characteristics as a basis for analyzing our qualitative (i.e., interview) data for students' meanings for independence and covariance.

Methods

Data Collection

Student participants were enrolled in an undergraduate Probability course for engineering and computer science majors in Spring 2024. For each participant we collected previous GPA and whether they had previously taken or were currently enrolled in Discrete Mathematics. We also gave students a survey asking about their exposure to logic instruction. We were prompted to do this by an interview participant in Phase 2; she had studied logic for LSAT preparation. The survey asked about other courses students had taken and other non-course related activities they had experienced. Through discussions with course instructors, we determined Discrete Math would be the only course used as an indicator for prior logic training, but students with select non-course experiences would be considered as having formal logic training.

The first author taught a 50-minute lecture to 7 sections of this course. This was a modification from our data collection in Phase 1, when only 4 sections were included in the study and were taught by two different instructors (Ryals et al., 2023). Moreover, in this phase of our study, all 7 sections were taught with the same lesson plan, similar to the standard treatment used in the prior phase. Student written work from this class meeting as well as responses to one midterm and one final exam problem were collected for all students present. There were 137 students who granted consent to participate, of whom 91 had received prior training in logic.

During the lecture, the instructor presented the theorem and discussed the possibility of the converse being true. Due to results obtained from interviews in Spring 2023 (Ryals et al., 2024), the instructor asked students for counterexamples of the converse. That is, she asked students how two variables could have a covariance of 0 but be dependent. In each of the 7 sections, at least one example was presented and discussed and this was used to clarify that the theorem is not biconditional. The instructor then worked two examples that asked first whether two variables were independent and then asked for their covariance. In one example, the variables were independent and therefore the theorem could be used to conclude the covariance was 0. In the other example, the variables were dependent, so covariance had to be calculated. Students were then presented with two problems. They worked these two problems individually and their work was collected. In both problems, the students were first asked to compute the covariance of two random variables and then were asked to find whether the variables were independent. In the first, which we refer to as the "converse problem," the covariance was 0, so students were then expected to use the definition of independence. In the second, which we refer to as the "contrapositive problem," the covariance was not 0, so students had the option of using the theorem to conclude the variables were not independent. We also collected student responses to one problem on a midterm assessment and one problem from the final exam. The midterm problem was very similar to the converse problem and the final exam problem was very similar to the contrapositive problem. Responses collected are characterized in Table 1 below.

Table 1: Sample Characteristics

	No Prior Logic Training	Prior Logic Training	Overall
Average Previous GPA	3.5	3.7	3.6
Number of Post-Class Work Responses	46	91	137
Number of Midterm Responses	35	61	96
Number of Final Responses	37	62	99

Analysis

We began by coding the responses from the in-class problems. Each of the three authors individually coded each participant's response to each of the two problems, primarily using the codes established in Phase 1 of the study (Ryals et al., 2023). We met to discuss and resolve the codes we applied individually. The modification of codes was guided by analysis from the interviews completed in Phase 2 of the study and from disagreements in codes selected by the different authors. For example, on the converse problem where covariance was 0, multiple students stated "X and Y are independent because E[XY]=E[X]E[Y]." In phase 1, we had coded this response as "Assumes the converse is true." However, analysis of one interview revealed that one student believed that the expectations being equal, as in the equation above, was the definition of independence. (She was not referencing the theorem.) Therefore, we created the code, "Used expectation or variance equations." After completing the coding for the in-class problems, two authors coded each student response to the midterm problem and two authors coded each response to the final exam problem. We again met to discuss and resolve any disagreements.

Statistical analyses were conducted to investigate the relationship between prior logic training and response type as well as to examine how student responses changed over time. Multinomial logistic regression was used to study the relationship between prior logic training and response type with prior GPA and prior logic training as covariates and response code as the dependent variable. McNemar-Bowker tests (Bowker, 1948) were conducted to study how student responses changed over time by comparing responses in the post-class problems to the corresponding problems on the midterm and final. Planned analyses included McNemar-Bowker tests, by logic training status, on the change in response from one converse problem, post-class problem 1, to the next, the midterm problem, as well as on the change in response from one contrapositive problem, post-class problem 2, to the next, the final problem.

Results

Types of Responses

We begin by presenting and summarizing the codes developed in Phase 3 which were modified from analysis in Phase 1 (Ryals et al., 2023). We will explain how these were condensed to quantitatively detect differences in how students interpret the variants of the theorem. Response rates for codes for each two types of problems are presented. Then we discuss inferential statistics assessing whether prior logic training impacted student responses and whether those responses changed over time.

Student responses closely resembled those in Phase 1 of the study. Each codes indicates how students responded to the question about independence and most often can be categorized as used the definition, used the value of covariance, used multiple arguments, or did not answer or provide justification.

Table 2: Phase 3 Codes Developed for Written Student Responses

Converse Problem	Contrapositive Problem
assumed converse was true	bypassed theorem (used definition)
claims independence can't be determined from covariance of 0	claims covariance not equal to 0 is not sufficient evidence to determine independence
used expectation or variance equations	used expectation or variance equations
no answer	no answer
incorrect or incomplete argument	incorrect or incomplete argument
no justification provided	no justification provided
used definition	used theorem
used conditional definition/argument	used theorem and also referenced definition
used informal definition argument	used definition and also referenced theorem
used or stated definition incorrectly	used informal definition/argument
wrong covariance calculation	Wrong covariance calculation
	bypassed theorem incorrectly

Prior to conducting analyses, the response codes listed in Table 2 were condensed into similar response types. For example, “used informal definition argument” and “used conditional definition/argument were combined with “used definition.” These condensed response types for each type of problem (converse, contrapositive) and their relative frequencies are listed below in Tables 3 and 4, respectively.

For the converse problems, relative frequencies stayed similar between the post-class problems and the midterm, though fewer students made the mistake of claiming that independence could not be determined from a covariance of 0. On the contrapositive problems, however, a large portion of students who used the theorem on the post-class problem reverted to using the definition of independence on the final. Both methods can be used to correctly answer the question, though using the theorem is more efficient.

Table 3: Relative Frequencies of Condensed Codes for Converse Problems

Code Number		Post-Class #1	Midterm
1	Used the definition or meaning of independence	59.1%	58.1%
2	Claimed independence cannot be determined from a covariance of 0	10.9%	3.2%
3	Used covariance of 0 or expectation or variance equations	26.3%	32.3%
4	No complete argument given	3.6%	6.5%

Table 4

Relative Frequencies of Condensed Codes for Contrapositive Problems

Code Number		Post-Class #2	Final
1	Used the definition or meaning of independence	13.1%	41.9%
2	Used theorem	81.0%	53.8%
3	No complete argument given	5.8%	4.3%

Multinomial logistic regression analyses were conducted on the student response codes for each question controlling for prior logic training and previous GPA. The results for post-class Question 1 are found below in Table 5.

Table 5: Coefficients for Multinomial Logistic Regression for Post-Class Question 1

Code Number	Intercept	Prior Logic Training	Previous GPA
2	2.71 p = 0.709	0.89 p = 0.838	0.49 p = 0.34
3	104.44 p = 0.013	1.71 p = 0.254	0.20 p = 0.002
4	35.97 p = 0.207	3.61 p = 0.292	0.13 p = 0.015

The intercept coefficients represent the odds of each student response code occurring compared to the reference code (1) when prior logic training and previous GPA are 0. As a previous GPA of 0 is not a reasonable possibility, these values are not interpreted. These results do suggest that prior logic training is not a significant indicator of any student response code at the 0.05 level. The results do, however, suggest that previous GPA is a significant predictor of student response code. The odds for student response code 3, compared to student response code 1, decrease by 80% for a one unit increase in previous GPA while the odds for student response code 4, compared to student response code 1, decrease by 87% for a one unit increase in previous GPA.

Multinomial logistic regressions were also conducted for post-class question 2 and the midterm and final questions. The results mirror those of the test for post-class question 1 with no statistically significant results at the 0.05 level for prior logic training. The smallest p-value found for prior logic training was 0.19. For the midterm question, which was a converse problem similar to post-class question 1, GPA was again a significant predictor of student response code. For this analysis, the odds for student response code 3, compared to student response code 1, decrease by 96% for a one unit increase in previous GPA while the odds for student response code 4, compared to student response code 1, decrease by 99% for a one unit increase in previous GPA.

To examine how student responses changed over time, McNemar-Bowker tests were conducted between post-class question 1 and the midterm question and between post-class question 2 and the final exam question. This test examines if there's a pattern in how students changed their answers from one question to the next. For example, imagine comparing two sets of answers from the same students. If about the same number of students changed their answers in each direction, the test would say there's no significant pattern. But if more students changed their answers in one direction than the other, the test would say there's a significant pattern.

Initially, tests by logic training status were planned, however, due to a smaller than expected number of students without prior logic training, conducting these tests by logic training status was not possible due to many missing response combinations. As a result, the McNemar-Bowker tests were conducted without regard to logic training status. For the comparison of post-class question 1 student response codes and the midterm question student response codes, the global test of symmetry has a p-value of 1, indicating that there is no pattern of movement from one student response code to another response code. For the comparison of post-class question 2 student response codes and the final question student response codes, the global test of symmetry has a p-value of 0.002, which is significant at the 0.05 level. This indicates a pattern of movement from one response code to another. To discover which pair or pairs of response codes have this pattern of moment, additional tests were conducted between each pair of response codes. The comparison of response codes 1 and 2 has a p-value of 0.007, when adjusted for the multiple comparisons. This indicates that there is a significant pattern of movement from the student response code of "Used theorem" on post-class question 2 to "Used the definition or meaning of independence" on the final. The comparison of response codes 1 and 3 has an adjusted p-value of 0.201 and the comparison of response codes 2 and 3 has an adjusted p-value of 0.617, indicating that there is no significant pattern of movement between codes 1 and 3 or codes 2 and 3.

Conclusion & Discussion

In this study we report on the third phase of a study investigating how students interpret variants of a theorem in probability relating covariance and independence of random variables. Using results from prior phases, we modified and applied an existing coding scheme and classified student written responses to two types of problems for 137 participants.

Upwards of 25% of students conclude that variables must be dependent when covariance is 0, immediately after instruction as well as weeks later on the final exam. This is despite an intended improvement to instruction in Phase 3, specifically presenting and illustrating a counterexample to the converse. These results support prior work showing students' tendencies to interpret conditional statements as biconditional and their approaches are not necessarily influenced by counterexamples (Sierpińska, 1987).

A large majority of students used the contrapositive of the theorem correctly in class when presented with two variables whose covariance is not 0. However, fewer students used the theorem on a similar problem on the final exam and this difference was significantly significant. We note that there was also an increase in the number of students correctly using the definition of independence on the converse problem on the midterm, in comparison to in class. Therefore, overall, it seems students become more comfortable using the definition of independence over time for both types of problems.

Consistent with results in Phase 1 (Ryals et al., 2023) and prior studies outside Probability (Attridge et. al, 2016; Cheng et. al, 1986; Inglis & Simpson, 2008), the differences in categories of responses for students with logic training and those without are not statistically significant. This is despite the fact that we more carefully considered how to classify a student's logic background in Phase 3.

We know from the quantitative results from this study that students' logic training may not be the only important factor to consider when discussing students' interpretations of this theorem. While logic training did not have statistically significant differences, GPA was a predictor for student success with this particular theorem. This has significant teaching implications because one might expect that some students with an overall lower GPA might have more difficulty interpreting the theorem as a conditional, one-directional statement. The quantitative data also led us to believe that there may be lexical ambiguities with i*ndependence*, *correlation*, and *covariance*, which might also be important factors to consider as we enter the next phase of this study. We further suspect that lexical ambiguities might be of concern because as we coded this data, we noticed that multiple students claimed that since Cov(X,Y)=0, X and Y are *uncorrelated*, and therefore are independent. We hypothesize that students' experience with colloquial uses of the words uncorrelated and independent have led them to view these words as synonyms. However, this needs to be further explored with student interviews that we recently conducted with this same pool of students for our next phase of the study. Future interviews may also be helpful in determining why some students provided multiple justifications, and why some used expectation or variance equations in lieu of covariance values.

References

Attridge, N., Aberdein, A., & Inglis, M. (2016). Does studying logic improve logical reasoning? 2016. In Csíkos, C., Rausch, A., & Szitányi, J. (Eds.). Proceedings of the 40th Conference of the International Group for the Psychology of Mathematics Education, Vol. 2, pp. 27–34. Szeged, Hungary: PME

Batanero, C., Chernoff, E., Engel, J., Lee, H., & Sanchez, E. (2016). Research on Teaching and Learning Probability. In: Research on Teaching and Learning Probability. ICME-13 Topical Surveys. Springer, Cham. https://doi.org/10.1007/978-3-319-31625-3_1

Batanero, C., Godino, J. D., & Roa, R. (2004). Training teachers to teach probability. *Journal of Statistics Education,* 12. https://jse.amstat.org/v12n1/batanero.html

Bowker, A. H. (1948). *A test for symmetry in contingency tables*. Journal of the American Statistical Association, 43(244), 572–574.

Case, J. (2015). Calculus Students' Understanding of Logical Implication and its Relationship to their Understanding of Calculus Theorems. University of Maine. Electronic Theses and Dissertations. Paper 2254.

Casella, G. & Berger, R. L. (2002). Statistical inference (2nd ed.). Duxbury

Cheng, P. W., Holyoak, K. J., Nisbett, R. E. & Oliver, L. M. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. Cognitive Psychology, 18, 293-328.

Dawkins, P. C. & Cook, J. P. (2017). Guiding reinvention of conventional tools of mathematical logic: students' reasoning about mathematical disjunctions. *Educational Studies in Mathematics, 94*(3), 241-256.

Durand-Guerrier, V. (2003). Which notion of implication is the right one? From logical considerations to a didactic perspective, *Educational Studies in Mathematics, 53*, 5-34.

Durst, S., & Kaschner, S. R. (2020). Logical Misconceptions with Conditional Statements in Early Undergraduate Mathematics. *PRIMUS*, *30*(4), 367-378.

Glasersfeld, E. v. (1995). Constructing agents: The self and others. In L. P. Steffe & P. W. Thompson (Eds.), *Radical constructivism: A way of knowing and learning* (pp. 113 – 128). New York, NY: RoutledgeFalmer.

Hoyles, C., & Küchemann, D. (2002). Students' understandings of logical implication. *Educational Studies in Mathematics, 51*(3), 193-223.

Inglis, M., & Simpson, A. (2008). Conditional inference and advanced mathematical study. *Educational Studies in Mathematics*, *67*(3), 187-204

Kaplan, J., Rogness, N., & Fisher, D. (2014). Exploiting lexical ambiguity to help students understand the meaning of r*andom*. *Statistics Education Research Journal, 13*(1). https://doi.org/10.52041/serj.v13i1.296

Kaplan, J. J., Fisher, D., & Rogness, N. (2010). Lexical ambiguity in statistics: How students use and define the words: association, average, confidence, random and spread. *Journal of Statistics Education*, *18*(2), 1–22. http://www.amstat.org/publications/jse/v18n2/kaplan.pdf

Kolmogorov, A. (1950). *Foundations of probability's calculation.* New York: Chelsea Publishing Company (Original work, published in 1933).

O'Brien, T. C., Shapiro, B. J., & Reali, N. C. (1971). Logical thinking-language and context. *Educational Studies in Mathematics*, 201-219.

Pfannkuch M. & Constance, M. B. (2017). Building on and Challenging Students' Intuitions About Probability: Can We Improve Undergraduate Learning? *Journal of Statistics Education,* https://doi.org/10.1080/10691898.1996.11910502

Piaget, J. (1977). *Research on reflective abstraction*. Paris: P.U.F.

Ryals, M., Hellings, C., & Ma, H. (2023). Students' inferences using a conditional statement in probability: The relationship between independence and covariance of random variables. In S. Cook, B. Katz, & D. Moore-Russo (Eds.), *Proceedings of the 25th Conference on Research in Undergraduate Mathematics Education.* (pp. 702-710). Omaha, NE

Ryals, M. & Sellers, M. (2024). Student interpretations of conditional statements in probability:

The cases of Byron and Sophie. In S. Cook, B. Katz, & D. Moore-Russo (Eds.), *Proceedings of the 26th Conference on Research in Undergraduate Mathematics Education.* (pp. 1133-1139). Omaha, NE

Sellers, M. (2020). *Students' quantifications, interpretations, and negations of complex mathematical statements from calculus* [Unpublished doctoral dissertation]. Arizona State University. https://keep.lib.asu.edu/items/158397

Sellers, M. E., Roh, K. H., & Parr, E. D. (2021). Student quantifications as meanings for Quantified variables in complex mathematical statements. *The Journal of Mathematical Behavior, 61,* 1-16. https://doi.org/10.1016/j.jmathb.2020.100802

Sellers, M. E. (2024). The problems with "any" and "all": How language impacts student meanings for quantified variables. In S. Cook, B. Katz, & D. Moore-Russo (Eds.), *Proceedings of the 26th Conference on Research in Undergraduate Mathematics Education.* (pp. 608-616). Omaha, NE

Sierpińska, A. (1987). Humanities students and epistemological obstacles related to limits. *Educational studies in Mathematics*, *18*(4), 371-397.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: techniques and procedures for developing grounded theory, 2nd edition.* Thousand Oaks, CA: Sage.

Thompson, P. W., Carlson, M. P., Byerley, C., & Hatfield, N. (2014). Schemes for thinking with magnitudes: An hypothesis about foundational reasoning abilities in algebra. In K. C. Moore, L. P. Steffe & L. L. Hatfield (Eds.), *Epistemic algebra students: Emerging models of students' algebraic knowing*, WISDOMe Monographs (Vol. 4, pp. 1-24). Laramie, WY: University of Wyoming. http://bit.ly/1aNquwz.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology, 20*(3), 273–281.