Textbook Formations of Independence

Steven Edalgo
Oklahoma State University

Adam Molnar
Oklahoma State University

The noun independence and adjective independent are applied in multiple mathematical contexts. In probability, independent events do not affect each other, but in algebra and regression, an independent variable has a non-symmetric effect on a dependent variable. Further complicating matters, independence in everyday language represents something in between. Prior research has shown that students and professors struggle to apply concepts of independence. As part of an investigation into curriculum about independence, textbook definitions about independence were examined. Across nine books, a mix of algebra and statistics texts, substantial variations existed in definitions of independent events and independent variables. Variations included the register of representation, verbal against algebraic, and the strength of the dependent effect. Little written guidance was provided to help learners navigate across the multiple formations.

Key words: Independence; Probability; Variable; Semiotics; Lexical Ambiguity

Independence is an old concept in probability. Over 250 years ago, De Moivre defined independent events in *The Doctrine of Chances*; two events are *independent* when "the happening of one neither forwards nor obstructs the happening of the other" and *dependent* if the "probability of either's happening is altered by the happening of the other" (1756, p. 6). The concept is symmetric; either event can serve as the "one" or the "other". In probability, the term maintains that definition today. Independent events are sufficiently common that the authors of the Common Core State Standards chose to include the definition in the high school standards (National Governors Association [NGA] Center for Best Practices & Council of Chief State School Officers [CCSSO], 2010, p. 82).

Despite 250 years of history—or perhaps because of 250 years of history—people have trouble determining if events are independent. Manage and Scariano (2010) surveyed 219 college students in US introductory statistics classes; only 23% correctly answered a multiple-choice question on the definition of independence. Molnar (2016) surveyed 25 US high school mathematics teachers; only 3 (12%) correctly solved a problem about two events in a table. Outside the USA, D'Amelio (2009) wrote about students' and professors' challenges in Argentina; about half the French pre-service teachers surveyed by Nabbout-Cheiban (2016) incorrectly solved problems about independent events.

One potential reason behind the trouble is that the colloquial non-probabilistic definition of independence differs from De Moivre's statement. According to the Oxford English Dictionary, the adjective *independent* refers to something "not depending on the authority of another, not in a position of subordination or subjection; not subject to external control or rule; self-governing, autonomous, free" ("Independent," 2015). In everyday language, independence and dependence are not necessarily symmetric. For instance, when choosing where to live, young children are usually dependent on their parents' decisions, but parents have more autonomy.

Another complication arises from the labeling of *independent* and *dependent* variables when describing algebraic functions. In Common Core standards, Grade 6 students should "write an equation to express one quantity, thought of as the dependent variable, in terms of the other quantity, thought of as the independent variable" (NGA Center for Best Practices & CCSSO, 2010, p. 44). With defined sides, this definition is never symmetric, although closer to the sometimes-symmetric everyday definition than the always-symmetric probability version. Incomplete textbook explanations do not assist students. Leatham (2012) gathered 10 school mathematics textbooks containing problems on independent and dependent variables. Of 73 total problems, 32 (44%) provided absolutely no information to determine the independent variable; others provided only partial information. Leatham concluded that most textbook problems sent mixed messages, "implicitly impeding students from developing a robust understanding of independent and dependent variables" (p. 357).

Some statistics textbooks tag on an additional non-symmetric definition of independent variables in regression models, where the independent variable controls the value of the dependent variable. Because the regression definition is similar to the algebraic one, and the definition does not appear in state standards, it is less crucial. Besides, between colloquial, algebraic, and probabilistic definitions, ample opportunities exist for confusion.

Our entire investigation will consider the curriculum around independence through multiple lenses as defined in Gerhke, Knapp, and Sirotnik (1992). In this preliminary report, we examine planned textbook curriculum in algebra and statistics textbooks. Later, we will interview teachers to ask about intended curriculum and analyze student artifacts of experienced curriculum.

Theoretical Framework

Independence is a non-physical mathematical concept. As Duval wrote in 2006, mathematical objects are never physically visible. Humans comprehend mathematics only through symbols and signs, "but the mathematical objects must never be confused with the semiotic representations that are used" (p. 107). Therefore, our framework for understanding mathematical confusions about objects labeled independent is semiotic.

For example, a mathematical object that can change between more than one value, dependent on circumstances, receives the representation *variable* in written English. In college algebra, the object receives a single letter representation such as X. In either setting, the object could have been expressed with another semiotic representation, such as changing-number or ϖ or 변수 (the Korean representation for the concept of variable).

In Duval's (2006) framework, systems of representations are called *registers*; earlier, variable was presented in the registers of written English, algebra, and Korean. Explanations for mathematical objects, known as *formations*, are designed to help students connect a concept with its representation in a register. Students are expected to learn formations. For instance, in the Common Core probability standards, students are asked to explain the formation of "independence in everyday language and everyday situations" (NGA Center for Best Practices & CCSSO, 2010, p. 82).

When authors write textbooks, they generate written formations for mathematical objects. Some authors include formations in other registers or ask students to convert between registers. Our textbook research uses the semiotic framework to ask the following questions.

1. For independent events, what formations and registers are used in definitions?
2. For independent variables, what formations and registers are used in definitions?

We had also proposed a third question, about explanations provided to distinguish applications of the semiotic symbol *independence*, but we did not find many of these explanations. We comment more about this in the section on questions for the audience.

Method

Encyclopedic search of pre-algebra, algebra, and probability textbooks would be very long, given the abundance of available textbooks. For this RUME report, we decided to concentrate on a small sample of books we knew had recent college-level use, similar to how Cook and Stewart (2014) examined recently-published textbooks on linear algebra. We selected three college introductory statistics textbooks (Bluman, 2014; Bowerman, O'Connell, & Murphee, 2014; Diez, Barr, & Cetinkaya-Rundel, 2015) and three college algebra textbooks (Bittinger, Beecher, Ellenbogen, & Penna, 2013; Crauder, Evans, & Noell, 2014; Miller, 2014). For comparative purposes, we added three US secondary school algebra textbooks (Benson, Dodge, Dodge, Hamberg, Milauskas, & Rukin, 1991; Bittinger, 1999; Brown, Dolciani, Sorgenfrey, and Kane, 1990). In each textbook, we recorded the initial formative definition involving *independent* or *independence* and examined problems in the text related to independence.

Results

Each introductory statistics textbook had a different definition of independent in regards to events. One college algebra textbook and one secondary school algebra textbook also contained formations because the books included sections on probability. The initial definitions are presented in Table 1, with statistics book definitions first.

Table 1
Initial Textbook Definitions Related to Independent Events in Probability

Textbook	Definition
Bluman (2014, p. 213)	Independent Events – Two events A and B are independent events if the fact that A occurs does not affect the probability of B occurring.
Bowerman et al. (2014, p. 171)	Independent Events – Two events A and B are independent if and only if 1.) $P(A\|B) = P(A)$ or, equivalently, 2.) $P(B\|A) = P(B)$ Here we assume that P(A) and P(B) are greater than 0.
Diez et al. (2015, p. 85)	Independent Process – Two processes are independent if knowing the outcome of one provides no useful information about the outcome of the other.
Brown et al. (1990, p. 756)	Independent Events – Two events A and B are independent if and only if: $P(A \cap B) = P(A)P(B)$.
Miller (2014, p. 779)	Independent Events – If events A and B are independent events, then probability that both A and B will occur is P(A and B) = P(A)*P(B).

The authors of a statistics textbook, Bowerman et al. (2014), presented the most challenging definition, using both the probability algebra register $P(A|B) = P(A)$ and conditional probability. The two algebra textbooks also contain formulations in the probability algebra register, varying slightly in the semiotic sign for and (and versus $\cap$), but do not require another mathematical concept. Relying on an additional concept complicates the structure. If a student cannot convert $P(A|B)$ into a mental concept, the student will not comprehend independence. Research results, summarized by Falk in 1986, have shown that conditional probability confuses many students.

Asking learners to construct a mental formulation of independence through another challenging concept, plus a conversion from the algebra register, is highly demanding.

On the other hand, Diez et al. (2015) and Bluman (2014) avoided conditional probability and the algebra register, relying only on written English. The computational formula appears later. Placing the written definition first reduces cognitive load by requiring less symbolic conversion. Nevertheless, despite language similarities, their definitions are not alike. A process is a larger concept than an event; events are sets of outcomes inside random processes. By defining independence on processes, not two events inside a process, Diez et al. (2015) have offered a different conception than the other authors. Interestingly, Diez et al. later refer to independent events, writing "if two events are independent, then knowing the outcome of one should provide no information about the other" (2015, p. 94). The shift between larger processes and smaller events may not appear notable, but for a concept with demonstrated problems, all shifts in formulation matter. Bluman's (2014) definition is the clearest.

Turning to variables, four algebra books and two statistics books contained a definition for independent variables. We do not know why the other two algebra books did not; perhaps the authors considered the concept a prerequisite. Initial definitions are presented in Table 2, with college algebra textbooks first, then secondary school algebra textbooks, then statistics books.

Table 2

Initial Textbook Definitions Related to Independent Variables

Textbook	Definition
Bittenger et al. (2013, p. 62)	Independent Variable – In the equation $y = \frac{3}{5}x + 2$, the value of y depends on the value chosen for x, so x is said to be the independent variable.
Miller (2014, p. 17)	Independent Variables – One type of mathematical model is a formula that approximates the value of one variable based on one or more independent variables.
Benson et al. (1991, p. 346)	Independent Variable – In a function, the variable whose value is subject to choice. The independent variable affects the value of the dependent variable.
Bittenger (1999, p. 156)	Independent Variables – Write an equation like $y = x^2 – 5$, which we have graphed in this section, it is understood that y is the dependent variable and x is the independent variable, since y is calculated after first choosing x and y is expressed in terms of x.
Bluman (2014, p. 19)	Independent Variable – In an experimental study, the one that is being manipulated by the researcher. … The resultant variable is called the dependent variable.
Bowerman et al. (2014, p. 487)	Independent Variable – Regression analysis is a statistical technique in which we use observed data to relate a variable of interest, which is called the dependent (or response) variable, to one or more independent (or predictor) variables.

As with the independent event definitions, we see multiple registers in Table 2. Both books with Bittenger (1999, 2013) as an author initially used the algebraic register. Although the letter x is a common algebraic formulation for an independent variable, writing in two registers complicates the concept. In other surveyed books, the symbolic x appears later.

Both Bittenger books (1999, 2013) introduce independence through the verb choose; independent variables have values selected, but dependent variables do not. The noun choice appears in Benson et al.'s (1991) definition; Bluman's (2014) definition of manipulation in experiments has synonymous language. The other two books do not mention choosing a value for independent variables. Both Miller (2014) and Bowerman et al. (2014) utilize a word related to prediction. Only Miller's 2014 mathematics book—not a probability and statistics text—describes the connection as approximate, although *affect* in Benson et al.'s (1991) definition is not as strictly causative as the other books or the Common Core formulation. Overall, free choice versus control appears to be the dominant formulation. The two statistics books that use the term independent variable do not vary much from algebra formulations, a slightly surprising result.

Questions for the Audience

The third probability and statistics textbook made a type of distinction we had hoped to see frequently. When discussing regression modeling, Diez et al. add in a footnote that applying the words independent and dependent "becomes confusing since a *pair* of variables might be independent or dependent, so we avoid this language" (2015, p. 18, emphasis in original). The other books do not make distinctions or connections between formations involving the word independent. One possibility for the paucity of connections we saw is our sample size. Although our recollections of other texts include few connecting and distinguishing statements, a larger search could identify more. Alternatively, we could interacting with more teachers and students.
1. Would a more comprehensive textbook search be fruitful?

The three concepts are unalike in the direction of the relationship. As described earlier, the probability relationship is symmetric; the algebra relationship is not symmetric; the colloquial relationship is sometimes symmetric. In an earlier discussion, one person suggested that relationships be defined in terms of causal direction, bi-directional versus uni-directional.
2. What research could be done to investigate this idea? Given the untested nature of any new causal definitions, the investigation would likely have to occur outside standard classroom flow.

This leads into another question, the primary one in the minds of teachers—and us.
3. What are other possible solutions to the misconceptions and lexical confusion?

References

Benson, J., Dodge, S., Dodge, W., Hamberg, C., Milauskas, G., & Rukin, R. (1991). *Algebra 1: An integrated approach*. Evanston, IL: McDougal Littell.

Bittinger, M. L. (1999). *Intermediate algebra* (8th ed.). Boston, MA: Addison Wesley Longman.

Bittinger, M. L., Beecher, J. A., Ellenbogen, D. J., & Penna, J. A. (2013). *College algebra: Graphs and models* (5th ed.). New York, NY: Pearson Education.

Bluman, A. G. (2014). *Elementary statistics: A step-by-step approach* (9th ed.). New York, NY: McGraw-Hill Education.

Bowerman, B., O'Connell, R., & Murphee, E. (2014). *Business statistics in practice* (7th ed.). New York, NY: McGraw-Hill Education.

Brown, R. G., Dolciani, M. P., Sorgenfrey, R. H., & Kane, R. B. (1990). *Algebra and Trigonometry*. Boston, MA: Houghton Mifflin Harcourt.

Carder, B., Evans, B., & Noell, A. (2014). *Functions and change: A modeling approach to college algebra* (5th ed.). Boston, MA: Houghton Mifflin Harcourt.

Cook, J. P., & Stewart, S. (2014). Presentation of matrix multiplication in introductory linear algebra textbooks. In Fukawa-Connolly, T., Karakok, G., Keene, K., & Zandieh, M. (Eds.), *Proceedings of the 17th Annual Conference on Research in Undergraduate Mathematics Education* (pp. 70–76). Denver, CO: Special

D'Amelio, A. (2009). Undergraduate student difficulties with independent and mutually exclusive events concepts. *The Montana Mathematics Enthusiast, 6*(1&2), 47–56.

De Moivre, A. (1756). *The doctrine of chances* (3rd ed.). London, England: Millar.

Diez, D. M., Barr, C. D., & Cetinkaya-Rundel, M. (2015). *OpenIntro Statistics* (3rd ed.). Boston, MA: Openintro, Inc. Retrieved from https://www.openintro.org/stat/textbook.php

Duval, R. (2006). A cognitive analysis of problems of comprehension in a learning of mathematics. Educational Studies in Mathematics, 61, 103–131.

Falk, R. (1986). Conditional probabilities: insights and difficulties. In R. Davidson & J. Swift (Eds.), *Proceedings of the Second International Conference on Teaching Statistics* (pp. 292–297). Victoria, Canada: International Statistical Institute.

Independent. (2015). In *Oxford English dictionary* (2nd ed.). Retrieved from http://www.oed.com/view/Entry/94325

Gehrke, N. J., Knapp, M. S., & Sirotnik, K. A. (1992). In search of the school curriculum. *Review of Research in Education, 18*, 51-110.

Leatham, K. R. (2012). Problems identifying independent and dependent variables. *School Science and Mathematics, 112,* 349–358.

Manage, A. B. W., & Scariano, S. M., A classroom note on: Student misconceptions regarding probabilistic independence vs. mutual exclusivity. *Mathematics and Computer Education, 44*(1), 14–21.

Miller, J. (2014). *College algebra* (1st ed.). New York, NY: McGraw-Hill Education.

Molnar, A. (2016, July). *High school mathematics teachers' understanding of independent events.* Paper presented at the 13th International Congress on Mathematical Education, Hamburg, Germany. Retrieved from http://iase-web.org/documents/papers/icme13/ICME13_L9_Molnar.pdf

Nabbout-Cheiban, M. (2016). Intuitive thinking and misconceptions of independent events: A case study of US and French pre-service teachers. *International Journal of Research in Undergraduate Mathematics Education.* Advance online publication. doi:10.1007/s40753-016-0038-x

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards: Mathematics.* Washington, DC: Authors.