*The Mathematics Educator*
2019 Vol. 28, No. 1, 3–26

# A Preliminary Genetic Decomposition of Probabilistic Independence

Karen Zwanch

*The purpose of this research is to construct a preliminary genetic decomposition delineating the mental constructions underlying probabilistic independence. This delineation is considered within the framework of APOS theory. While the use of the term independence in probability is often conflated with causation, the definition relies instead upon an understanding of conditional probability. I hypothesize that the concept of independence is only fully available to students by constructing at least an object conception of probability. I offer additional hypotheses, supported by literature and anecdotal teaching experience, regarding students' quantification of probability and construction of combinatorial reasoning.*

Many researchers have documented the overwhelming difficulty students experience in trying to construct appropriate understandings of probabilistic tasks (Jones, Langrall, Thornton, & Mogill, 1997; Tarr & Jones, 1997). In particular, the concept of independence and dependence of events is problematic. Shaughnessy (2003) indicates that "students have difficulty just sorting out the mathematics of whether events are statistically dependent or independent in probability problems" (p. 221). Many researchers (D'Amelio, 2009; Nabbout-Cheiban, 2017; Ollerton, 2015) agree that this is due to a disconnect between an intuitive sense of the terms as they are generally used in the English language and the mathematically correct definitions.

The conditional definition for independence is at times problematic for students as it occasionally works in contrast to the students' colloquial sense of the word. The definition is as follows: Two events, $A$ and $B$, are independent if and only if the occurrence of event $B$ has no effect on the probability of the occurrence of event $A$. That is, $P(A) = P(A|B)$ and $P(B) = P(B|A)$ (Grinstead & Snell, 1997). Dependent events are

***Karen Zwanch*** *is an assistant professor of mathematics education at Oklahoma State University. Her primary research focuses on using middle-grades students' concept of number to model their algebraic reasoning.*

defined as events that are not independent; two events, $A$ and $B$, are dependent if and only if $P(A) \neq P(A|B)$.

Students' abilities to correctly interpret this definition have implications for their use of probabilistic formulas in a variety of situations (Ollerton, 2015). One such situation is that if two events, $A$ and $B$, are independent, then $P(A \cap B) = P(A) \cdot P(B)$. This theorem results from the conditional definition (Grinstead & Snell, 1997), as demonstrated below.

$$P(A \cap B) = P(A) \cdot P(B|A) \tag{1}$$

This is true for all events, regardless of independence

$$P(B) = P(B|A) \tag{2}$$

Definition of independent events

$$P(A \cap B) = P(A) \cdot P(B) \tag{3}$$

Result of substituting (2) into (1)

The resulting theorem in Equation 3 is mathematically preferred in some situations and can be used to test for independence. Its use in instruction to replace the conditional definition, however, may be a confounding factor in students' difficulties with probabilistic independence by redirecting focus from conceptual understanding to completing calculations.

One particular difficulty students have with independence is a reliance on their intuitive sense of the concept. D'amelio (2009) found that undergraduates persistently apply their intuitions regarding independence. Forty-six students out of 54 in her research study responded to the question of two events being independent intuitively by indicating, for example, that the events had no connection. Similarly, Nabbout-Cheiban (2017) found that for preservice teachers, intuition and colloquial interpretations of the term independence were common and problematic. For instance, preservice teachers commonly considered independent events to be a "one-way relation of one event to another" (p. 273), rather than to be a property describing the relationship between the two events. This is attributed to the application of a colloquial definition of independence because of the assumption that one event is not dependent on the other, for instance. Furthermore, when the preservice teachers solved problems involving independent events, they frequently gave intuitive responses in place of completing calculations.

If students' intuitive ideas about independence are disconnected from formal definitions or calculations, they will not likely develop conceptual insight, especially if taught Equation 3 without understanding its derivation. In this study, I examine the possibility of constructing the underlying conceptual structures necessary to accept the conditional definition before introducing the concept of independence. As a means of determining what the necessary conceptual structures might be, I will construct a preliminary genetic decomposition of the concept of independence of events.

## Theoretical Framework

The theory supporting the genetic decomposition includes APOS theory and Piagetian stages.

### APOS Theory

APOS Theory delineates the process by which individuals mentally construct mathematical schemas by constructing action (A), process (P), and object (O) conceptions of mathematical concepts (Dubinsky, 1991), the coordination of which formulates schemas (S). Schemas are "structures that contain the descriptions, organization, and exemplifications of the mental structures that an individual has constructed regarding a mathematical concept" (Arnon et al., 2014, p. 25). Schemas are dependent upon the mathematical ideas that an individual perceives to be interrelated. The transformations lead to more sophisticated conceptions and are explained by reflective abstraction (Piaget, 1977/2001), which entails reflection on concepts or operations being applied to the concept. Reflection leads to the advancement of the concept to a higher cognitive level that constitutes, from Piaget's perspective, the learning of mathematical ideas, and in APOS theory, a more sophisticated conception.

An action conception is the most rudimentary. With an action conception, completing mathematical tasks is procedural

(Arnon et al., 2014), and it is necessary for students to carry out the actions of solving because mental re-presentations are not yet available. Reflecting on actions leads to the interiorization of a process, with which students can assign meaning to mental re-presentations of concepts. A process conception allows students to reflect upon, describe, and reverse mental processes (Asiala et al., 1996). Reflecting on mental processes can lead to the encapsulation of a process into an object, which implies that conceptual objects are entities in and of themselves that have properties, can be considered in totality, and to which actions can be applied (Asiala et al., 1996). Students' conceptions are organized in a schema, which can itself be a conceptual object to which actions are applied; this enables less sophisticated schemas to be included within more sophisticated schemas (Arnon et al., 2014), creating a cyclical process by which interconnected schemas are constructed.

Students' applications of actions, processes, and objects are not strictly linear. Once actions are interiorized into processes, processes can be encapsulated into objects to which actions can be applied. However, objects can also be de-encapsulated into processes (Arnon et al., 2014). Thus, a student who has encapsulated a mathematical concept into an object can act in a manner that is consistent with having constructed only a process, for example.

## Piagetian Stages

The research of Piaget and Inhelder (1951/1975) regarding students' constructions of chance and probability is also included within this model. Piaget and Inhelder delineated students' reasoning through three developmental stages. In stage I, the preoperational stage, children may use symbols but do not mentally manipulate objects or information; this aligns with Dubinsky's (1991) actions, in which problem solving is procedural. Stage II is the concrete operational stage and is marked by the onset of inductive logic and reversibility. Reversibility, in particular, is also a marker for a process in APOS (Asiala et al., 1996). The final stage, stage III, is formal operational and supports abstract thought, thereby aligning approximately with objects in APOS (Asiala et al., 1996). Although APOS theory and Piaget's stages are distinct from one another, their similar epistemological foundations suggest compatibility. Piaget and Inhelder's research contributes empirical findings that are relevant, and will therefore be integrated into this research to build a more robust, and integrated genetic decomposition.

## Genetic Decomposition

Within the frame of APOS theory, "a new mathematical concept frequently arises as a transformation of an existing concept" (Arnon et al., 2014, p. 28). Thus, a genetic decomposition is a trajectory of concepts and the transformations of those concepts that result in the construction of new or more sophisticated concepts. Arnon et al. (2014) state that a genetic decomposition "explains whatever is known about students' expected performances that indicate differences in the development of students' constructions. … [and] include[s] a description of prerequisite structures an individual needs to have constructed previously" (p. 28). They further explain that designing a genetic decomposition can include self-reflection, reflection on work with students, or extant literature on student thinking. Finally, while a genetic decomposition may inform instructional decisions, it is not a teaching sequence.

This study is a preliminary genetic decomposition of probabilistic independence; it is preliminary in the sense that it has yet to be empirically tested and refined (Arnon et al., 2014). Accordingly, analysis included reflections on the concept of independence, reflections on related teaching experiences, and existing literature on student thinking. The resulting model includes mental constructs that support the construction of independence, transformations that occur to these mental constructs in service of constructing independence, and descriptions of student performances that may provide indications of the existence of such mental constructs. I created the model through a comparison of reflections and literature aimed at detailing the constructs that support a sophisticated concept of probabilistic independence.

Some analytic choices guided the construction of this specific model. The model culminates in the conceptualization of the conditional definition of probabilistic independence. That is not to discount other possible conceptualizations, only to suggest that the conditional definition is a basis upon which other conceptualizations of independence can be built. This model also focuses on discrete as opposed to continuous sample spaces. While the resulting model could apply to continuous sample spaces, probabilistic independence involving continuous spaces does not require combinatorial reasoning. This model incorporated combinatorial reasoning for discrete sample spaces so as to build a more comprehensive model. Genetic decompositions are not unique (Arnon et al., 2014), which implies the possibility of other models, and other models may include a more narrow or wide scope than what is presently defined. The model constructed in this research, which is based on a systematic comparison of literature and reflections on probabilistic independence, resulted in a hypothetical model for students' construction of the conditional definition of probabilistic independence.

## Conceptual Analysis

Thompson (2008) describes a conceptual analysis as relevant to "describing ways of knowing that might be propitious for students' mathematical learning, and … that might be deleterious to students' understanding of important ideas" (p. 46). Both students' favorable and unfavorable ways of understanding probabilistic independence will be central to the construction of this genetic decomposition. To illustrate the conceptual analysis, I will focus on a problem of tossing three dice. The problem has two parts: (a) If three dice are tossed simultaneously, what is the probability of rolling a triple? (b) If three dice are tossed in succession, what is the probability of rolling a triple? Actually, these questions are equivalent and both represent independent events; however, many students view them as being different and may interpret the second question to be an example of dependent events. As a result of the belief that the events in the second situation are dependent, they may view the result as having either a higher or a lower probability than that of the first situation.

Students relying on a colloquial interpretation of independence may indicate that the events in part b of the dice problem are dependent because after the first die is rolled, the number that must result on the second and third dice is set because the outcome of the second and third dice generating a triple is dependent on it being the same as the outcome of the first die. As the number that each die must land on is now predetermined, students may further engage in subjective reasoning to determine that the probability of the second and third dice landing on the same number as the first is less likely than was the probability of the first die rolling any number. According to Jones, Langrall, and Mooney (2007), this type of causal or deterministic reasoning is a common cause of misconceptions regarding probability.

In contrast, applying the conditional definition of independence leads to the conclusion that rolling the dice simultaneously or successively are statistically equivalent. Consider, for instance, rolling a three on the first and second dice as events $A$ and $B$, respectively. The probability of rolling three on the first die is 1/6. The probability of rolling three on the second die, given the result of three on the first die, remains 1/6. Therefore, $P(B|A) = P(B)$ and thus events $A$ and $B$ are independent; the same result follows for the independence of the third roll. Because the simultaneous, and successive, rolls of three dice are independent, calculating the probability of rolling any set of triples can be done using the equation resulting from the definition of independence of events (Equation 3). In other words, $P(n \cap n \cap n) = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{216}$, where $n$ represents an arbitrary integer on the die. Therefore, the probability of rolling triples is the sum of the probabilities of rolling any specific triple: $P(triples) = 6\left(\frac{1}{216}\right) = \frac{1}{36}$.

The solution to this problem is demanding as it requires students to consider many different probabilistic components simultaneously, the first of these being the independence of events. Without an appropriate conceptualization of independence, students tend to rely on subjective reasoning which creates confounding errors throughout their solution. Both simple and conditional probabilities are inherent within the conditional definition of independence. As conditional probability is necessarily dependent on simple probability, I will begin with the mental constructs necessary to conceptualize simple probability. I will then build to an understanding of conditional probability and, finally, independence of events.

## Mental Constructs Underlying Probability

Simple probability is the calculation of the chance of the occurrence of a single event, which is represented as a ratio of favorable outcomes to total outcomes. Jones et al. (2007), in their synthesis of probability research, characterize probability in the following way: "*Probability* focuses directly on describing, quantifying, modeling, and illuminating random processes" (p. 910). In following with this characterization, students must accomplish significant mental constructions to engage in simple probabilistic reasoning. To begin, I will outline how understanding randomness, quantifying probability, and modeling probability contribute to simple probability. Following this outline, I will also consider how simple probability fits within the APOS framework with regard to probability as a whole.

Randomness is one of the mental constructs that contribute to simple probability. Piaget and Inhelder (1951/1975) illustrate children's construction of randomness by having them predict and explain the mixing of marbles in a box. At first children may anticipate that all marbles will return to their original position despite mixing. Children with more sophisticated reasoning explain that as the box is tilted more, the marbles will become increasingly and irreversibly mixed; however, these children often explain random mixtures to emulate a pattern. Thus, it is not until children demonstrate understanding that the arrangement of marbles does not emulate a pattern that they have constructed randomness.

In addition to constructing randomness, children must construct the quantification of probability (Jones et al., 2007), which relies on their construction of the logical operation of disjunction (Piaget & Inhelder, 1951/1975). Piaget and Inhelder (1951/1975) used the construct of disjunction to describe the child's reasoning about the composition of a set. Consider, for instance, presenting a child with two bags of marbles. In the first bag are one red and one blue marble. In the second bag are two red and three blue marbles. Although the probability of selecting a red marble from the first bag is higher (1/2 compared to 2/5), a child who has not constructed disjunctive reasoning may indicate it is more likely to draw a red marble from the second bag because there are two reds. This is because the child is considering only the favorable outcomes (one red marble compared to two red marbles) and ignoring the composition of the set as a group of favorable and unfavorable outcomes combined (one red out of two marbles compared to two red out of five marbles). In summary, a child who has not yet constructed disjunctive reasoning may be unable to compare the favorable and unfavorable outcomes within a set to determine which is more likely. These children instead focus only on the favorable outcomes, ignoring the unfavorable outcomes altogether.

In their theoretical framework to understand young children's quantification of probability, Jones et al. (1997, 1999) echo the importance Piaget and Inhelder (1951/1975) place on disjunctive reasoning by indicating that part-whole and part-part relationships are of great relevance in conceptualizing probability. Furthermore, they view children's development of part-whole and part-part reasoning to be a critical factor in students' probabilistic growth. Within Jones et al.'s (1997) framework, students progress through four levels of reasoning with regard to probability: (a) subjective, (b) transitional, (c) informal quantitative, and (d) numerical. Although the authors of this neo-Piagetian framework do not explicitly state the construction of disjunctive reasoning to be a predicating factor in a student's construction of probability, they label the first two levels in such a way as to imply its importance as an underlying mental construct. In level one, judgments about probability are subjective, and in level two favorable and total outcomes are not considered in tandem (Jones et al., 1997). Both of these characterizations point to students not having constructed disjunctive reasoning.

The final component that Jones et al. (2007) identify is modeling probability. This includes the classical, or theoretical, and the frequentist, or experimental, view-points. The classical perspective involves quantifying the expectation that an event will occur whereas the frequentist perspective involves quantifying the occurrence of an event obtained through data collection (Shaughnessy, 1992). Although the frequentist perspective acts in support of the classical perspective, it necessarily represents a truncated quantification of the probability of an event. To construct an understanding of the relationship between these two perspectives requires students to construct the law of large numbers.[1] Furthermore, Piaget and Inhelder (1951/1975) indicate that students do not include the law of large numbers appropriately in their reasoning until they have fully constructed randomness. Thus, Jones et al.'s (2007) final component of understanding of probability, modeling probability, is based on the reconciliation of classical and frequentist views of probability. This reconciliation is made possible by the law of large numbers which depends on the construction of randomness.

## Sample Space

An additional consideration in modeling probability is sample space, which is the ability to determine the entire set of possible outcomes (Jones et al., 1997), and in simple probability, the outcomes are those from one event. With only one event, students' difficulties with sample space are generally related to their subjective decision making (Jones et al., 1999), as opposed to difficulties constructing a list of outcomes. For example, while a child making subjective judgments about sample space likely understands that a single die can land on any integer between one and six, the child is unlikely to understand the implications of this sample space to be that any integer has a one

[1] Students' construction of the law of large numbers is a discussion beyond the scope of this research.

in six chance of being rolled. This is representative of the child's disinclination to connect sample space to probability (Jones et al., 2007), and aligns with Piaget and Inhelder's (1951/1975) indication that children cannot reason about chance until they have constructed disjunctive reasoning because the logical operation of disjunction supports reasoning about the set of favorable and possible outcomes. Thus, when considering simple probability, the disconnect between sample space and probability is more so the cause for difficulty than the construction of the sample space itself.

As probabilistic situations become more complex, however, the construction of sample space becomes more complex as well. The construction of sample space for multiple events involves various forms of combinatorial reasoning, including combinations and permutations. I argue that the ability to manipulate and combine combinations and permutations requires a schema that coordinates both. The following section delineates the constructions of such schemas.

**Permutations.** When students consider multi-stage experiments, they must also consider permutations as a means of constructing sample space. Permutations may be more difficult for students to construct, in comparison to combinations, because students must first come to understand that AB is distinct from BA (Piaget & Inhelder, 1951/1975), which is irrelevant with combinations. Permutations will be considered first, however, because the derivation of the formula for permutations is necessary for the formula for combinations.

After beginning to distinguish between permutations containing the same elements, such as AB and BA, children can construct an action conception of permutations. In the early stages of conceptualizing permutations, Piaget and Inhelder (1951/1975) find that students rely on physical models or manipulatives, which is due to limitations of their mental structures. I will discuss students' construction of permutations within APOS theory in the context of the four-color cubes problem, in which students are asked to make all two-cube permutations from the set of four, without replacement.[2]

With a rudimentary concept of permutations, children will likely use some sort of manipulatives or models to determine all the possible outcomes of an experiment. Piaget and Inhelder (1951/1975) find that children at stage IA do not initially understand AB and BA, for example, to be distinct permutations. As they come to understand different arrangements of the same pairs of cubes constitute different permutations, which is indicative of the construction of stage IB, they are likely to “grope” with manipulatives on a problem like the color cubes problem (Piaget & Inhelder, 1951/1975). This groping constitutes an action conception (Asiala et al., 1996) of permutations because children do not have a systematic means by which they form permutations and they have no way to determine if they have found all permutations. Children might arbitrarily make and record groups of two color cubes, but they may make the same group again without realizing, or they may not find all the groups. With only an action conception of permutations, pairs of cubes are individually counted and children do not anticipate them before they are formed. Thus, students are likely to rely on manipulatives or other models, and the arrangements of these manipulatives or models may seem arbitrary because they are created in the moment. The odometer strategy (English, 1991), and models such as area and tree diagrams, enable children to organize their permutations. Each of these strategies provides a systematic means by which the color cubes can be arranged exhaustively.

Construction of a conception beyond an action results from reflecting on these models. Supported by manipulatives and models of permutations, students can begin to generalize patterns. Regarding the example of the color cubes problem, students will notice that when the first color cube (cube A) is

[2] I refer to the color cubes throughout my discussion of permutations and combinations. This example is adapted from the work of Piaget and Inhelder (1951/1975). By color cubes, I mean four small cubes, each of which is a different color. For clarity, I will refer to them as cubes A, B, C, and D.

placed first, three groups of two cubes can be formed (AB, AC, AD); when the second color cube (B) is placed first, there are three more groups (BA, BC, BD); and so on. As students reflect on this pattern, they interiorize the actions of finding permutations into a process (Asiala et al., 1996), which is evidenced by students beginning to anticipate the number of possible combinations and to run through the activity mentally. These reflections align with Piaget and Inhelder's (1951/1975) description of children at Stage IIB who observe regularities in permutations, have intuitive anticipations, and develop a "progressive consciousness of symmetries" (p. 187). These descriptions align with a process conception of permutations because students begin to anticipate and mentally run through creating permutations. Thus, with a process conception, the model and manipulatives can be used to *verify* that one's set of groups is exhaustive. This verification is a nuanced, but important, distinction from the manner by which students with only an action conception of permutations rely on models and manipulatives to *generate* all the groups.

With a process conception, students can reflect on their mental processes further, thereby engendering the encapsulation of the mental process into an object. Regarding permutations, students can generalize that the first selection in a permutation comes from a group of $n$, the second from a group of $n-1$, and the number of permutations when selecting two objects is $n \cdot (n-1)$. This aligns with Piaget and Inhelder's (1951/1975) descriptions of students at stage IIIA, for whom systems begin to emerge and multiplicative reasoning comes into play. For example, on the color cubes problem, students might begin to reason that there are 12 permutations because there are four cubes to select from for the first element of the set and three cubes remaining to select from for the second element of the set $(4 \times 3 = 12)$.

Additionally, with an object conception, students can compose permutations, which enables them to mentally construct *sets* of permutations in which more than two selections are being made. For example, a permutation of three events requires permutations on two selections and then an operation on that preliminary set of permutations. Composing permutations facilitates the generalization of the pattern that additional selections increase the number of permutations to $n \cdot (n-1) \cdot (n-2)$, and so on, and eventually to $n!$. Composing multiple permutations in this way typifies an object conception. This generalization is distinguished by Piaget and Inhelder (1951/1975) as available to students only at stage IIIB, indicating it to be the most advanced development of students' reasoning about permutations. Within APOS theory, when a schema includes "a coherent collection of structures . . . and connections established among those structures, it [schema] can be transformed into a static structure (Object)" (Arnon et al., 2014, p. 25). In terms of permutations, this implies that the schema containing students' actions, processes, and objects of permutations can itself be encapsulated into an object upon which students can act by composing permutations, thereby generalizing the formula *n!*.

**Combinations.** Combinations are also necessary for students to reason about sample space in multi-stage experiments. Like permutations, a child's conception will begin as an action and will require manipulatives or models to support their reasoning. Consider the color cubes problem again. Children at Piaget and Inhelder's (1951/1975) stage I will likely "grope" with the cubes because they do not anticipate the combinations that result from the activity of selecting two cubes. Organizing their results using models or the odometer strategy (English, 1991) provides a systematic means by which the color cubes can be arranged, as well as a visualization of repeated or impossible combinations. Interiorizing these actions into a process results from reflections on these actions.

Supported by models of combinations, students can begin to generalize patterns; these generalizations are typical of Piaget and Inhelder's (1951/1975) stage II. In the color cubes problem, for example, students notice that by placing the first color cube (cube A) first, three groups of two cubes can be formed (groups AB, AC, and AD); when the second color cube (B) is placed first, there are two more groups (BC and BD); and so on. As students reflect on this pattern, they can interiorize the actions of finding combinations into a mental process (Asiala et al., 1996); namely, they can begin to anticipate the number of possible combinations and run through the activity mentally. Thus, with a process conception, models or manipulatives can be used to verify that all groups have been found rather than to find all of the groups as a student with only an action conception would do. This allows students who have constructed a process conception of combinations to justify that their set of groups is exhaustive.

Encapsulating a process for combinations into an object is characterized by students' abilities to begin conceiving of the set of combinations in the color cubes problem, for example, without manipulatives or models. Piaget and Inhelder (1951/1975) note that children in stage III are able to determine combinations without any physical manipulatives because they can coordinate several sets of pairings simultaneously. For example, a student at stage III can coordinate the goal of pairing block A first with the other three blocks (i.e., AB, AC, AD) while pairing block B first with the other two blocks (i.e., BC, BD), and so on. Piaget and Inhelder's characterization of students at stage III aligns with having encapsulated a process conception of combinations into an object because students are operating on sets of combinations.

**Combinatorial reasoning.** As discussed previously, a schema for permutations can include students' generalization of the formula *n!* once they encapsulate permutations into an object. A schema including combinations and permutations, which I will refer to as combinatorial reasoning, allows for the joint consideration of combinations and permutations in constructing sample space. Furthermore, deriving a formula for combinations is more complex than that of permutations in that it requires students to coordinate schemas for combinations and permutations. Thus, the justification of this formula is considered within a student's combinatorial reasoning schema rather than as existing within a student's combinations schema. The formula for combinations without replacement is $\binom{n}{r} = \frac{n!}{r!(n-r)!}$. It involves three components, each of which is related to a conceptual idea: that combinations are a specialized case of permutations ($n!$), that duplicated cases are eliminated through division by *r!*, and that selecting fewer than *n* items is accounted for through division by *(n-r)!*. Students must reflect upon and generalize these relationships in order to derive the formula. Thus, a conceptual understanding of a derivation of the formula for combinations is considered evidence of a student's construction of a schema of combinatorial reasoning that coordinates permutation and combination schemas. It is this concept of combinatorial reasoning that acts in support of advanced conceptions of probability. The delineation of probability within APOS theory will be considered next.

## Probability and Probabilistic Independence

Conditional probability is defined in the following way: "The conditional probability of an event *A*, given that event *B* has occurred, written $P(A|B)$, is the probability of *A* considering as possible outcomes only those outcomes of the random experiment that are elements of *B*" (Tarr & Jones, 1997, p. 40). This definition illuminates the difficulty of conditional probability over simple probability. Tarr and Jones (1997) note that students have particular difficulty constructing the appropriate sample space for conditional probability because they fail to reevaluate the sample space based on the conditioning event. Accordingly, I will delineate the mental constructions students must make in order to appropriately conceptualize conditional probability, and ultimately independence (Figure 1).

The construction of simple probability involves randomness, modeling probability, and the quantification of probability (Figure 1, arrows a, b, and c); furthermore, simple probability will begin as an action. Having constructed only an action, students cannot mentally anticipate results. As such, students are unlikely to discern a relationship between classical and frequentist results of probability because an action conception does not support the anticipation that the infinite extension of an experiment would generate the theoretical result. To interiorize an action into to a process requires reflection on the results of the classical and frequentist perspectives (Figure 1, arrow d). This reflection includes the generalization that the results of the two types of problems are equal, thereby indicating

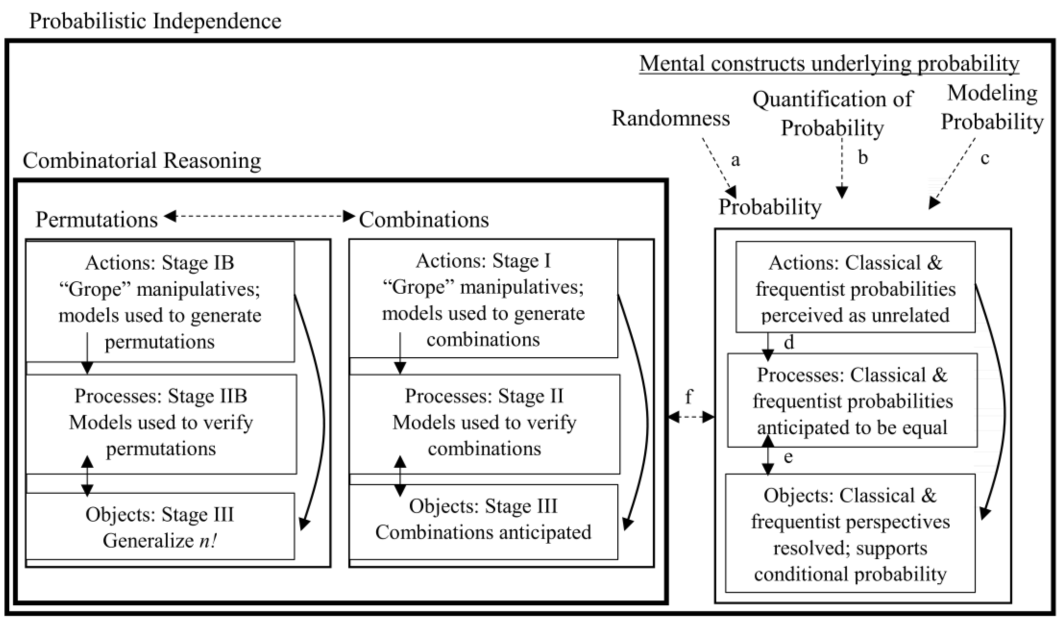


*Figure 1*. Preliminary genetic decomposition of probabilistic independence. Solid arrows indicate transformations of concepts among stages. Dashed arrows indicate coordinations among schemas. Stages are interpreted from the work of Piaget and Inhelder (1951/1975).

that the two actions (conducting an experiment and comparing favorable to all outcomes) are the same.

A process conception enables more advanced probabilistic reasoning, which includes comparing probabilistic results without actions. With a process conception, students can anticipate the results (Arnon et al., 2014) of a single theoretical event, facilitating comparison between the likeliness of two events. Having constructed a mental structure for probability also allows students to reverse the process of probability (Arnon et al., 2014), which allows students to create a set of all possible outcomes based on known probabilities. For instance, if the probability of drawing a red marble from a bag is $\frac{3}{10}$ and there are only red and blue marbles, a student with a process conception can determine that there may be three red and seven blue marbles in the bag. Students with a process conception of probability have reflected on the results of classical and frequentist perspectives to the extent that they expect them to generate the same result, and will likely be perturbed if they do not. Students who have constructed only a process may lose track of the idea that the equivalence of results relies on a sufficiently large number of trials because their mental structure for probability anticipates the similarities they have observed and reflected on over time.

Reflecting on mental processes leads to the encapsulation into objects (Figure 1, arrow e). Specifically, reflection on the distribution of results from experimental data and the manner in which experimental results approach theoretical probabilities supports construction of an object. Thus, the connection between theoretical and experimental results is resolved with an object conception, meaning that students now conceive of theoretical and experimental probabilities to represent different perspectives on the same probability. The strength of an object conception is that it enables students to conceptualize compound and conditional probabilities because an object implies that probability can be acted upon by combining multiple events (i.e., compound probability) and finding probabilities nested within probabilities (i.e., conditional probability; Arnon et al., 2014). Furthermore, students can now coordinate probability with combinatorial reasoning (Figure 1, arrow f) to conceptualize compound and conditional probabilities, and they have finally constructed the mathematics necessary to begin conceptualizing the independence of events. By definition, “Two events are *independent* if the occurrence of one does not change the probability of the occurrence of the other” (Tarr & Jones, 1997, p. 40). Conceptualizing probabilistic independence as it is defined conditionally requires a schema conception that includes probability as an object so that a nested conditional probability can be conceptualized and compared to a simple probability.

While the incorporation of conditional probability into a schema for independence is necessary, it is not sufficient. In APOS theory, a schema can itself become an object to which actions are applied (Arnon et al., 2014). Thus, with an object conception of probability, students can begin to apply the action of calculating independence using the conditional definition. This is, however, only a jumping off point for conceptualizing independence because this action is, at first, external. Through reflections on this action, it can be interiorized into a mental structure for anticipating the results of determining independence, which constitutes a process (Arnon et al., 2014). This anticipation manifests behaviorally as students’ intuitions. Through reflections on intuitions, a process can be encapsulated into an object, in which the independence schema is an entity to which further actions are applied (Arnon et al., 2014). Accordingly, with an object conception of independence, students can overcome intuitive reasoning regarding independence because they can reason about and act upon independence as its own entity.

## Implications

The purpose of this preliminary genetic decomposition is to outline the mental constructs necessary to support a conception of probabilistic independence, informed by literature and reflection. Research indicates that a limited understanding of independence is problematic for secondary and undergraduate students (D’Amelio, 2009; Nabbout-Cheiban, 2017; Ollerton, 2015; Shaughnessy, 2003). Some students struggle with probabilistic independence in spite of a sophisticated conception of probability (e.g., Kelly & Zwiers, 1988; Nabbout-Cheiban, 2017; Ollerton, 2015). More specifically, students often reason about independence intuitively and with biases (D'Amelio, 2009; Nabbout-Cheiban, 2017; Ollerton, 2015), or by applying an understanding of the term as it is used colloquially rather than as it is defined mathematically (Thompson & Rubenstein, 2000; Kaplan, Rogness, & Fisher, 2014).

The implication of the extant literature is that probability is a necessary but not a sufficient condition for conceptualizing probabilistic independence. This genetic decomposition provides a lens through which the research regarding students' reliance on intuition and colloquial definitions can be understood. In the absence of sufficiently advanced mental constructs regarding probabilistic independence (e.g., combinations, permutations), students may rely on intuitions and apply biases to understand independence. Interpreting students' difficulties with independence through the framing of this genetic decomposition suggests that the students in existing literature who apply biases have not constructed conceptions of probability that support the construction of an object conception of independence.

This model demonstrates that in order to conceptualize the conditional definition of independence, students must have not only sophisticated conceptions of probability and combinatorial reasoning, but must also encapsulate their conception of independence into an object. Furthermore, while the result of the conditional definition $P(A \cap B) = P(A) \cdot P(B)$ is a useful mathematical tool for determining independence, it is not necessarily appropriate for students who have yet to encapsulate independence into an object. Therefore, instruction should focus first on supporting students' construction of the conditional definition of independence before encouraging students to apply the resulting theorem in its place.

Keeler and Steinhorst (2001) call for the improvement of instruction of probabilistic independence by capitalizing on students' intuitions. They posit that inquiry-based learning environments facilitate understanding by building on students' intuitions. Abrahamson (2014) also encourages probability instruction that "guide[s] students to appropriate the cultural resource as a means of supporting and empowering their tacit inference" (p. 250). His findings indicate that students' intuitions are a powerful instructional tool that instructors can leverage by linking to formal mathematics. Relating these results to the present research, I hypothesize that reflection on the mental processes that manifest behaviorally as intuitions about independence can potentially support students' encapsulation of those processes into objects, and teachers can leverage students' intuitions to support such an encapsulation.

According to Nabbout-Cheiban (2017), students with limited conceptions of combinatorial reasoning and probability may be successful on independence tasks by applying intuitions. From my own teaching experience, I also observe that some students operate with limited conceptions of probability and are intermittently successful with independence due to supports. For example, providing students who cannot coordinate probability and combinatorial schemas with sample space in a two-way contingency table may support them in successfully determining the independence of two events. Alternatively, students with limited conceptions may be successful on complex problems by relying on formulas. However, these students may struggle to derive the formulas because deriving formulas requires constructing probability and combinatorial schemas that are sophisticated enough to support independence. Therefore, these types of difficulties may be helpful indicators to instructors whose students need more time and support to construct sophisticated conceptions of independence. These students may initially require supports, such as a two-way table or support applying formulas, to be successful. However, it is more important to provide these students with opportunities to reflect on these support materials in order to engender their construction of more sophisticated conceptions of independence. The empirical refinement of this genetic decomposition could test this hypothesis.

It is easy to mistake students' intermittent success with understanding, and so it is necessary to take a deeper look at the limitations of students' reasoning. A more comprehensive understanding of how students construct the concept of probabilistic independence, and the underlying mathematical concepts, is necessary to advance research and teaching practices in this area. This preliminary model supports future research into how students construct the concept of probabilistic independence and provides preliminary evidence that is useful to instructors.

## References


Abrahamson, D. (2014). Rethinking probability education: Perceptual judgment as epistemic resource. In E. J. Chernoff & B. Sriraman (Eds.), *Probabilistic thinking: Presenting plural perspectives* (pp. 239–260). New York, NY: Springer.

Arnon, I., Cottrill, J., Dubinsky, E., Oktac, A., Fuentes, S. R., Trigueros, M., & Weller, K. (2014). *APOS theory: A framework for research and curriculum development in mathematics education*. New York, NY: Springer.

Asiala, M., Brown, A., DeVries, D. J., Dubinsky, E., Mathews, D., & Thomas, K. (1996). A framework for research and curriculum development in undergraduate mathematics education. In A. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Research in collegiate mathematics education II* (pp. 1–32). Washington, DC: Mathematical Association of America.

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

Dubinsky, E. (1991). Constructive aspects of reflective abstraction in advanced mathematics. In L. P. Steffe (Ed.), *Epistemological foundations of mathematical experience* (pp. 160–202). New York, NY: Springer-Verlag.

English, L. D. (1991). Young children's combinatorics strategies. *Educational Studies in Mathematics, 22*, 451–474.

Grinstead, C. M., & Snell, J. L. (1997). *Introduction to probability* (2nd ed.). United States of America: American Mathematical Society.

Jones, G.A., Langrall, C.W., & Mooney, E.S. (2007). Research in probability: Responding to classroom realities. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 909–956). Reston, VA: National Council of Teachers of Mathematics.

Jones, G. A., Langrall, C. W., Thornton, C. A., & Mogill, A. T. (1997). A framework for assessing and nurturing young children's thinking in probability. *Educational Studies in Mathematics, 32*(2), 101–125.

Jones, G.A., Langrall, C.W., Thornton, C.A. & Mogill, A.T. (1999). Students' probabilistic thinking in instruction. *Journal for Research in Mathematics Education, 30*(5), 487–519.

Kaplan, J. J., Rogness, N. T., Fisher, D. G. (2014). Exploiting lexical ambiguity to help students understand the meaning of *random*. *Statistics Education Research Journal, 13*(1), 9–24.

Keeler, C., & Steinhorst, K. (2001). A new approach to learning probability in the first statistics course. *Journal of Statistics Education, 9*(3). Retrieved from www.amstat.org/publications/JSE/v9n3/keeler.html

Kelly, I. W., & Zwiers, F. W. (1988). Mutually exclusive and independence: Unraveling basic misconceptions in probability theory. In R. Davidson & J. Swift (Eds.), *Proceedings of the Second International Conference on Teaching Statistics* (pp. 1–26). Victoria, VC: University of British Columbia.

Nabbout-Cheiban, M. (2017). 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, 3*, 255–282.

Ollerton, R. L. (2015). A unifying framework for teaching probability event types. *International Journal of Mathematics Education in Science and Technology, 46*(5), 790–794.

Piaget, J. (2001). *Studies in reflecting abstraction.* (R. L. Campbell, Trans.). New York, NY: Psychology Press. (Original work published 1977)

Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children.* (L. Leake, Jr., P. Burrell & H. D. Fischbein, Trans.). New York, NY: Norton. (Original work published 1951)

Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 465–494). New York, NY: Macmillan.

Shaughnessy, J.M. (2003). Research on students' understanding of probability. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 216–225). Reston, VA: National Council of Teachers of Mathematics.

Starnes, D. S., & Tabor, J. (2015). *The practice of statistics* (6th ed.). New York, NY: W. H. Freeman and Company.

Tarr, J. E., & Jones, G. A. (1997). A framework for assessing middle school students' thinking in conditional probability and independence. *Mathematics Education Research Journal, 9*(1), 39–59.

Thompson, D. R., & Rubenstein, R. N. (2000). Learning mathematics vocabulary: Potential pitfalls and instructional strategies. *The Mathematics Teacher*, *93*(7), 568–574.

Thompson, P. W. (2008). Conceptual analysis of mathematical ideas: Some spadework at the foundation of mathematics education. In O. Figueras, J. L. Cortina, S. Alatorre, T. Rojano, and A. Sepulveda (Eds.) Proceedings of the Joint Meeting of PME 32 and PME-NA XXX (Vol. 1, pp. 31 –50). Morelia, Mexico: Cinvestav-UMSNH.