STUDENTS' REASONING ABOUT THE ASSOCIATION OF CATEGORICAL VARIABLES USING CONTINGENCY TABLES AND MOSAIC PLOTS

by

SHERI ELLEN JOHNSON

(Under the Direction of Denise A. Spangler and Kevin C. Moore)

ABSTRACT

Researchers have called for more work related to students' understanding of two categorical variables. Additionally, it is widely agreed that reasoning across multiple representations can benefit conceptual understanding. In this study, I consider how middle and high school students reason about (in)dependence with contingency tables and mosaic plots. Clinical think-aloud interviews with middle and high school students reveal their efficient reasoning and productive use of a mosaic plot.

INDEX WORDS: association, categorical variables, contingency tables, mosaic plots, statistics, student reasoning

STUDENTS REASONING ABOUT THE ASSOCIATION OF CATEGORICAL VARIABLES USING CONTINGENCY TABLES AND MOSAIC PLOTS

by

SHERI ELLEN JOHNSON

B.I.E., Georgia Institute of Technology, 1988

M.S.H.S., Georgia Institute of Technology, 1989

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2020

STUDENTS' REASONING ABOUT THE ASSOCIATION OF CATEGORICAL VARIABLES USING CONTINGENCY TABLES AND MOSAIC PLOTS

by

SHERI ELLEN JOHNSON

Major Professors: Denise A. Spangler
Kevin C. Moore

Committee: Daniel B. Hall
Jori N. Hall

Electronic Version Approved:

Ron Wolcott
Interim Dean of the Graduate School
The University of Georgia
May 2020

DEDICATION

I dedicate this dissertation to my family, especially to my husband, Arlin Johnson, to my children Andrew, Robert, Daron and Elise Johnson, to my father, Robert Haas and to the memory of my mother, Lorraine Haas who each, in their unique way, have encouraged, supported, forgiven, accepted, and inspired me. Their unconditional love, unwavering faith, and joyful companionship are invaluable.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Denise Spangler who is an inspiring example, thoughtful leader, intelligent mentor, and accomplished researcher. I am blessed to consider her a critical friend. I would also like to thank my co-advisor, Kevin Moore for allowing me to sit in on his class years ago, which ignited my interest in mathematics education and quantitative reasoning. I appreciate his guidance throughout my dissertation study and his questions that pushed my thinking; I am a better researcher, interviewer, and writer because of him. I am grateful to Dan Hall for his statistical expertise, clarity of explanations, and high expectations. I am continually amazed by Jori Hall and appreciate her support, perspective, and qualitative research expertise. I also want to thank Kelly Edenfield for her unending availability and resources for any mathematics and statistics curriculum imaginable.

I am fortunate to have had an opportunity to learn what good research is from Andrew Izsák and Sybilla Beckmann and to have other colleagues on the sidelines with whom I could wonder and question (Merve, Rui, Torrey, and others). There have been so many other accomplished mathematics education faculty at UGA (Jeremy Kilpatrick, Jim Wilson, Les Steffe, Pat Wilson, Ryan Smith, Dorothy White, Jaime Diamond, and Anna Connor) and fellow graduate students (Merve, Kristen, Rui, Yuling, Kirsten, Jonathan, Halil, Biyao, Oguz, Eun, Eric, Burak, Claire, Julia, Irma, Pierre, Jernita, Dongjo, Dario, Brandon, Bolonle, and Dean) who have helped me on my journey. I am a better person and mathematician thanks to Ed Azoff and Zerotti Woods.

However, I would not have found my footing and another place to plant my work without Christine Franklin and the inspiring work she has done in statistics education. Thanks to Jennifer Kaplan for getting me started in statistics education research and to Kristen Roland for staying at UGA and mutually getting us through. Working with the GAISE II team (Christine Franklin, Anna Bargagliotti, Rob Gould, Pip Arnold, Leticia Perez and Denise Spangler) towards the end of this journey has been a gift for which I am most appreciative. I am hopeful that future generations will have the opportunity to become statistically literate citizens through their Pre-K–12 education.

Speaking of future generations, I am most thankful for my interview participants. These eight amazing middle and high school students who took the time after school, on weekends, and during school breaks to explain their thinking are the true basis of this study. My hope is that I have done some sort of justice to their time and thoughts and that it will make some positive impact somewhere, somehow.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

We live in a society where data are being captured and used in a variety of ways. Data can be used to document, explain, predict, and influence. “Big Data” is here, and to understand how it is being used, citizens should be statistically literate. Literacy, in all subjects, is an expected schooling result, and statistical literacy includes “the ability to interpret, critically evaluate, and communicate about statistical information and messages” (Gal, 2002, p. 1). This includes understanding the statistical mantra that association does not mean causation, often exemplified by comparing seasonal ice cream sales and drownings. There is an association between ice cream sales and drownings, but it is invalid to claim that eating ice cream causes drowning or vice versa without more evidence. Arguably more fundamental than the difference between association and causation is the understanding of association itself, or the difference between association and independence. Association can exist between variables that are *quantitative* (continuous or discrete), like a person’s height or weight, as well as variables that are *categorical* (nominal or ordinal), like a person’s eye color or hair color. Categorical variables are always discrete and can be created from partitioning quantitative variables (e.g., age cohorts). *Statistical association* refers to whether two variables are dependent (associated) or independent (not associated) based on results of statistical tests, and the two variables can be both quantitative, both categorical, or a mix of the two.

Categorical data are pervasive in surveys of opinions and also prevalent in health sciences, public health, physical sciences, education, marketing, and engineering (Agresti, 2007). Individuals deal with categorical variables daily with media and personal decisions – as much or more than quantitative variables. But, in teaching statistics we have traditionally spent more time with quantitative variables. It is important to understand the results of summative analyses. Understanding what it means for two categorical variables to be associated is necessary to become a statistically literate citizen. In efforts to create statistically literate citizens, there is now a substantial amount of statistics and probability in the K—12 curriculum. Statistics and probability content is mainly situated in the mathematics content of middle (23% of mathematics is probability and statistics) and secondary (20% of mathematics is probability and statistics) grades (Usiskin & Hall, 2015). Authors of the Common Core State Standards (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) stress the importance of reasoning with quantities in real-world problem contexts, and this includes beginning to reason about association.

Association of two categorical variables can be determined from data summarized in a contingency table, which is a tabular representation that displays the frequencies for all combinations of categories of variables in rows and columns. Late secondary students struggle to determine association with contingency tables (Batanero et al., 1996) because this requires coordinating multiple quantities and reasoning with ratios and proportions. Whereas ratios and proportions are a focus of middle grades mathematics, they are challenging concepts for students and adults (Beckmann & Izsák, 2015; Behr et al., 1992; Izsák & Jacobson, 2017; Kaput & West, 1994; Karplus et al., 1983a; Lamon, 2007;

Tourniaire & Pulos, 1985). To enhance proportional reasoning, many instructional materials for ratios and proportions include visual representations like double number lines, strip diagrams, and 100 blocks to assist students in making sense of this content (Beckmann, 2018; Van de Walle et al., 2019).

Research-based guiding principles include promoting reasoning as well as using and connecting mathematical representations (National Council of Teachers of Mathematics, 2014). However, in most classrooms in the US, "mathematics is presented as an almost entirely numeric and symbolic subject, with a multitude of missed opportunities to develop visual understanding" (Boaler et al., 2016, p. 1). Categorical association can be seen in other representations besides contingency tables including those that promote visual understanding. One example is a mosaic plot (see Figure 1), which is a representation that uses the concept of a unit square comprised of rectangular regions in a relative scale to depict the contingency table values. A mosaic plot is a potentially useful tool in assisting students to appropriately apply proportional reasoning to determine independence (Pfannkuch & Budgett, 2017).

Figure 1

Contingency Tables and Corresponding Mosaic Plots

A

		Hair color		Row Totals
		Light	Dark	
Eye color	Light	58	36	94
	Dark	13	47	60
Column Totals		71	83	154

B

		Eating habits		Row Totals
		Unhealthy	Healthy	
Recycle?	Yes	57	38	95
	No	36	24	60
Column Totals		93	62	155

Note. Panel A: Situation with association. Panel B: Situation with independence.

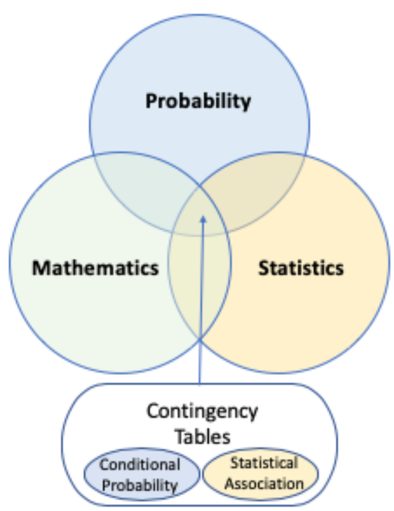
Reasoning across multiple representations is a tool students can use in problem solving and it deepens both procedural and conceptual understanding (National Council of Teachers of Mathematics, 2014). Using a mosaic plot in conjunction with a contingency table affords students this opportunity to reason across representations. Reasoning about association of categorical data is challenging content that requires proportional reasoning and it is necessary to develop this reasoning to become a statistically literate citizen. Possible association can be found numerically in a contingency table by either comparing or subtracting proportions (58/94 > 13/60 or 58/94 – 13/60 > 0, see Figure 1A). This possible association can also be seen in a mosaic plot by comparing the areas (the area of the light hair and light eyes as a proportion of the top row is greater than the area of the dark eyes and light hair as a proportion of the bottom row) or the horizontal distance (the horizontal length of the light eyes and light hair section is greater than the horizontal length of the dark eyes and light hair). Similarly, independence can be found numerically in a contingency table by either comparing or subtracting proportions (57/95 = 36/60 or 57/95 – 36/60 = 0, see Figure 1B). This independence can also be seen in a mosaic plot by comparing the areas (the area of the recycle: yes and unhealthy eating habits as a proportion of the top row is equal to the area of the recycle: no and unhealthy eating habits as a proportion of the bottom row) or the horizontal distance (the horizontal length of the light eyes and light hair section is equal to the horizontal length of the dark eyes and light hair).

Background

Reasoning with contingency tables lies at the intersection of three types of content: mathematics, statistics, and probability (see Figure 2). The language used with contingency tables differs based on whether they are used for probability or statistics (Watson & Callingham, 2014). With probability, language focuses on events, probabilities, and conditional probabilities, and it can support the understanding of Bayes Theorem. With statistics, language focuses on frequencies, relative frequencies, and joint, marginal, or conditional values, and it can support the understanding of association.

Figure 2

Contingency Tables Placed in Content



Extant research related to contingency tables considers either a focus on probability or a focus on statistics, and whereas my current interest of categorical association lies in statistical association, literature from both areas is pertinent because the underlying reasoning is similar. As a note, I use (in)dependent to denote situations in which it is not yet determined whether there is association (dependence) or independence between the variables. For instance, I use the phrase *determining (in)dependence* to refer to the goal of identifying whether or not two variables are associated.

Rationale

This study is crucial because understanding association with categorical data is an important component of being a literate citizen. With more surveys being conducted and presented through various forms of media, it is important that individuals can interpret, question, and critique data including how they are presented and the inferences that are being made. Statistics are increasingly available and used, so it is becoming more important for people to understand what statistics represent and “adopt a healthy questioning attitude towards what is presented”(Shaughnessy, 2007, p. 964). Data for categorical variables are also often analyzed in medical studies where the information informs real-life decisions. For example, someone might need to decide between different courses of treatment based on information that might consider the different status of disease (present and absent) for patients treated with different drug regimens. In these instances and others, understanding categorical data and reasoning to determine (in)dependence is necessary.

It is also important to identify how students make sense of graphical representations for the data that are summarized in contingency tables. Technology allows easy generation of graphs of various kinds, so it is important to develop a better understanding of how students make sense of these new forms of representations.

Research Questions

Students in late middle and early high school begin to reason about statistical association. In the context of statistical association, categorical data has unique nuances and challenges. Hence, it is important to understand how students reason with categorical variables and (in)dependence. Yet limited research has been done with categorical variables, and most of what has been done is with late or post-secondary students outside of the US. This research has focused on reasoning with completed contingency tables that are provided to students without consideration of other visual representations. Given the current emphasis on visual representations and making connections across representations, it is important to understand student reasoning in this regard. Thus, I am interested in examining the statistical thinking of students using both contingency tables and mosaic plots. Because proportional reasoning is a necessary complement for understanding (in)dependence in contingency tables, student participation was limited to those who showed evidence of robust proportional reasoning. I surmise that studying the reasoning of students related to contingency tables and mosaic plots is important, relevant, and valuable, and in this study I address the following research questions:

1. In what ways do students reason about (in)dependence of categorical variables when using contingency tables?
2. In what ways do students use mosaic plots to reason about (in)dependence of categorical variables when using contingency tables?

CHAPTER 2

LITERATURE REVIEW AND FRAMEWORK

Reasoning with contingency tables requires working with numbers in context. The context is defined by the categorical variables, and numerical reasoning includes both considering the whole numbers that represent the counts for categories of variables and reasoning about their relationships with one another. Thus, the numerical reasoning required includes fractions and ratios. In this chapter, I first review the pertinent literature and then based on the literature, suggest an appropriate framework.

Literature Review

Reasoning with Measured Quantities

Reasoning about categorical association requires more than just numerical reasoning; it involves reasoning with measured quantities while considering variability. Thompson (2010) identified that a quantity is comprised of multiple components including an object, an attribute of an object, and a unit of measure with an anticipated number that is proportional to this unit of measure. The units of measure do not have to be specified to reason with a quantity, but a quantity can be measured, resulting in a numerical value and a unit of measure (e.g., 3 inches). Reasoning with measured quantities is different than reasoning with quantities or numbers alone. For example, 3 + 2 + 6 = 11 requires only reasoning with numbers whereas reasoning with measured quantities might entail adding 3 apples, 2 oranges, and 6 kiwis together to result in 11 pieces of fruit. Quantitative operations are based on the situation and are not simply numerical calculations (with measured quantities), but numerical operations can evaluate a quantity (Thompson, 1994). For example, a quantitative operation for the fruit example will recognize that adding more fruit will result in a greater number of pieces of fruit, and it also involves coming to understand the process of creating a collective whole of fruit from the constituent collections of fruit types. The numerical operation of $3 + 2 = 5$ evaluates the total number of apples and oranges. The ability to follow a prescribed algorithm does not demonstrate quantitative reasoning. The quantities themselves and the relationships between them are created by an individual as they work through a problem (Moore, 2014), and what quantities and relationships a student knows should not be taken for granted (Izsák, 2003).

In statistics, "data are not just numbers, they are numbers with a context" (Cobb & Moore, 1997, p. 801). The raw data gathered for a contingency table are observations where the unit of measure is the category, and the numbers in a contingency table represent a statistical summary of a measured quantity as opposed to a number without context. For example, a population might be categorized by whether or not each person receives a drug and also whether or not they have a disease at some length of time after taking the drug. The observations are measured according to the presence or absence of the drug and the presence or absence of the disease. The data are then summarized to account for people who were given the drug and have the disease, people who were given the drug and do not have the disease, etc. These statistical summaries themselves are a measured quantity where the unit of measure is a count (cardinality).

Statistical Reasoning

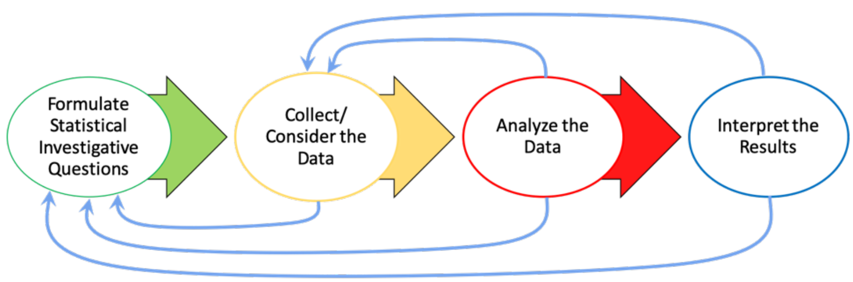
Reasoning about association with contingency tables requires context, and the two variables and their categories in a contingency table provide context. Statistics is different than mathematics because of the importance of context where "In mathematics, context obscures structure. In data analysis, context provides meaning." (Cobb & Moore, 1997, p. 803). Statistics is defined by the American Statistical Association as "the science of learning from data, and of measuring, controlling and communicating uncertainty." (Davidian & Louis, 2012, p. 12).

Variability can be quantified for both quantitative and categorical variables. A focus on variability and questioning the data are unique to statistics in comparison with other mathematical sciences (Bargagliotti et al., 2020). Answers in statistics often do not include a single, precise answer like they often do in mathematics. However, students should still be encouraged to employ the mathematical practice of "Attend to precision" (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010), and they can do this in a mathematical sense by calculating accurately and in a statistical sense by measuring variability and selecting appropriate representations that display variability (C. Franklin et al., 2006). Students should learn to appreciate variability and realize that because randomness exists, an answer is often a range of values rather than a single value.

Students should learn to do statistics by participating in the full investigative process (Van de Walle et al., 2019). This statistical problem-solving process (see Figure 3) is not necessarily linear, and students should use questioning throughout all components (Bargagliotti et al., 2020). A statistical investigative question is similar to a research question and may begin a study; however, when secondary data are used, a study might start by considering the data. Survey questions are used to gather data and analysis questions are used to interrogate the data. When students interpret the results, beginning students use descriptive statistics and as they progress in their statistical reasoning, they advance to drawing inferences from sample populations to general populations and between populations.

Figure 3

Statistical Problem-solving Process



Note. From *Pre-K–12 Guidelines for Assessment and Instruction in Statistics Education II (GAISE II)*, (p. 8), by A. Bargagliotti, C. Franklin, P. Arnold, R. Gould, S. Johnson, L. Perez, and D. Spangler, 2020, American Statistical Association. Copyright 2020 by the American Statistical Association.

Questioning is central to the statistical problem-solving process (Arnold, 2008), and continually questioning throughout is especially important. This is true for categorical data that might be presented in a contingency table and any other data display of summary statistics of raw data. Questions of interest include: How was the study designed? Where did the data come from? Did the data have to be cleaned or modified in some way? How can we best communicate what this data means?

Mathematical Reasoning

Reasoning about association with contingency tables requires proportional reasoning (Inhelder & Piaget, 1955, as cited in Batanero, Estepa, Godino, & Green, 1996), which in turn relies on reasoning with ratio and fractions and a sense of invariance (Lobato & Ellis, 2010). Fractions and ratios are challenging content; their definitions are often conflated, and their overlap is ambiguous.

Fractions and Ratios

Whereas students and teachers alike might view fractions and ratios as interchangeable, (Clark et al., 2003), most researchers agree they are different. For example, a ratio of concentrations of 3 different chemicals in a substance might be 2:5:7, and that is not a fraction. Also, a fractional measurement of a length may be ¾ of an inch or equivalently 0.75 inches, and that is a single number that can be placed on the number line. Whereas a fraction or single number can compare the magnitudes of a ratio of two numbers (e.g., 3/4 or 0.75 for a ratio of 3:4), there is not a single number for ratios with more than two numbers. A popular textbook designed for use with pre-service teachers formerly identified fractions as a subset of ratios (Van de Walle, 1994); however a more recent version (Van de Walle et al., 2018) and other sources claimed there is an overlap between fractions and ratios (Clark et al., 2003; Lobato & Ellis, 2010) yet they disagree as to what the overlap contains. Clark et al. (2003) argued that a part-part comparison can be a fraction whereas (Lobato & Ellis, 2010) considered a part-part comparison to only be a ratio, not a fraction, noting that "ratios can be meaningfully reinterpreted as fractions" (p. 27). To clarify this, they used an example of salad dressing with vinegar and oil in a ratio of 2:5 where the amount of vinegar is 2/5 the amount of oil. In this instance, 2/5 is not considering the whole of 7 parts of the mixture of salad dressing, but rather the whole has been redefined as the 5 parts of oil. The fraction still maintains a part-whole relationship, just with a different referent, the oil as opposed to the salad dressing.

"A ratio is a multiplicative comparison of two quantities, or it is a joining of two quantities in a composed unit" (Lobato & Ellis, 2010, p. 17). Additionally, a ratio might compare more than 2 quantities, as in the example of 1:2:3, but the aspect of multiplicative comparison remains where this can be thought of as 3 ratios 1:2, 2:3, and 1:3. Ratio reasoning starts in middle grades and compares two quantities multiplicatively, not additively. Ratios and multiplication both have additive entailments where the numbers in a part to part ratio can be added together to yield the collective whole (white and black paint in a ratio of 5:7 can be thought of as a combined mixture having 12 total parts, 5 white and 7 black) and multiplication can be thought of as repeated addition (3 x 5 can be thought of as 3 + 3 + 3 + 3 + 3 or 5 + 5 + 5). Younger students often compare quantities additively when it is appropriate to compare them multiplicatively (Noelting, 1980). Whereas an additive comparison answers the question "*how much more?",* a multiplicative comparison answers the question "*how many more times?"*. Consider the ratio of 4:2 where 4 might represent the number of apples and 2 the number of bananas. Multiplicative reasoning recognizes the number of apples is 2 times the number of bananas and equivalent ratios are 2:1, 6:3, 8:4, etc., whereas additive reasoning recognizes there are 2 more apples than bananas and using it results in non-equivalent ratios of 3:1, 5:3, 6:4, etc. Although some researchers supported "the model of multiplication is repeated addition" (Fischbein, Deri, Nello, & Marino, 1985, p. 6), other researchers mentioned it is an expired rule (Karp et al., 2015). Thinking of multiplication beyond just repeated addition is important in the development of proportional reasoning (Lobato & Ellis, 2010), and many students do not make this progression (Lamon, 2007). An extended meaning of multiplication considers equal-sized groups and the number of items in a group as a composed unit (e.g., 5 groups of 3) rather than adding 3 five times. This type of reasoning allows for the concept of scaling a number (e.g., 5 times the original value).

A fraction is defined as *a*/*b* where there are *a* copies each of size *1*/*b* (Van de Walle, Karp, Bay-Williams, & Wray, 2019). From a part-whole view of fractions, *1*/*b* can be thought of as a whole partitioned into *b* equal-size parts. Then iterating this part of size *1*/*b* exactly *a* times with no spaces between parts or overlapping parts results in a collective size of *a*/*b*. Another definition for fraction states: "*a*/*b* of our whole is the amount formed by *a* parts (or copies of parts), each of size *1*/*b* of the whole" (Beckmann, 2018, p. 48). Note these definitions do not preclude improper fractions, unlike a definition for a fraction of *a* "out of" *b* parts. These appropriate definitions also allow for multiplicative comparison, which is asking how many *b*'s are in *a* where the result is a numerical value.

Students struggle with fractions in various ways. More naïve reasoning includes thinking that a fraction is always less than 1 and that 1/5 is bigger than 1/3 because 5 is bigger than 3. Recent efforts surrounding fractions assist students to see a fraction not as two separate integers, but as a quotient, as a number with a value on the number line. Recognizing equivalent fractions where the numerical value is invariant is important to later develop proportional reasoning. Recognizing a fraction as a number also supports the ability to see a fraction as an operator, where it can be used to scale other values multiplicatively. This view of a fraction as an operator can be beneficial when working with contingency tables.

Despite efforts to improve proportional reasoning abilities, eighth-grade students continue to demonstrate difficulty correctly ordering fractions (Siegler, 2017). For example, students might be asked to determine which is greater, 3/4 or 5/7. Finding a common denominator allows comparisons of numerators to determine the greater value (e.g., comparing 3/4 to 5/7 with a common denominator of 28 allows 21 to be compared with 20 respectively). Using a decimal notation value of a fraction or standardizing ratios to percents can be helpful to compare amounts. Decimal notation of a fraction is the quotient of the numerator and denominator ($3/4 = 3 \div 4 = 0.75$), and a percent standardizes a ratio or fraction to have a denominator of 100 ($3/4 = 75/100 = 75\%$). The magnitude of proportions (in the statistical sense) can be represented as a fraction, decimal, or percent. When proportions are represented in decimal notation or as a percent, they can be compared using whole number reasoning (0.750 is greater than 0.714 or 75% is greater than 71.4% because 750 is greater than 714).

Probability

Fractions are a central topic in third through fifth grades, but probability, which is closely connected to fractions, is not addressed until seventh grade (Van de Walle et al., 2019). Probability measures the chance that an event occurs and can be represented on a number line from 0 (impossible) to 0.5 (equally likely) to 1 (certain). Probabilities can be theoretical or empirical. The theoretical probability is the ratio of the number of desired outcomes (event) to the number of all possible outcomes (sample space). For example, a fair die has 6 possible equiprobable results (1, 2, 3, 4, 5, 6), and the probability of rolling an even number (2, 4, 6) is 3/6 or 1/2. If the event is not in the sample space, the probability is 0 (the probability of rolling a 7 with one die is 0), and if the event is the only item in the sample space, the probability is 1 (the probability of rolling any number between 1 and 6 is 1). In some situations, the theoretical probability does not exist because the population is too large to accurately count or there is randomness in the events. The empirical probability is based on an experiment or observations, and it is the ratio of the number of times the desired outcome occurred (event) to the total number of outcomes. For example, rolling a die multiple times may result in 43 rolls of an even number in 100 total rolls where the empirical probability is 43/100. Contingency tables are generally used for results of experiments or observations, and thus the probabilities are generally empirical. However, if the total population is represented in the contingency table, then the relative frequencies are the same as the theoretical probabilities.

Conditional relative frequencies are akin to conditional probabilities, which are notoriously challenging (Kahneman & Tversky, 1982; Watson & Moritz, 2002). Conditional probabilities require redefining the sample space for theoretical probabilities or redefining the total number of observations for empirical probabilities. Similarly, from a part-whole perspective, conditional relative frequencies require redefining the whole from the total frequency to a marginal frequency. Contingency tables are used for probability and statistics alike, and there are many constituent components to coordinate when reasoning. When working with contingency tables, it can be confusing that marginal frequencies can be used as both the desired number of observations (the numerator) and the total number of observations (the denominator). Researchers demonstrated that multiple representations, including tree diagrams, can be helpful and some sixth-grade students were able to work with sample spaces (Nunes et al., 2014). However, not all students are given this same opportunity to work with multiple representations like tree diagrams.

Proportional Reasoning

Ultimately, reasoning about association with contingency tables requires proportional reasoning, which is a mature mathematical way of thinking. A proportion is defined differently in statistics and mathematics. The statistical use of proportion refers to the fraction of the total that has a certain attribute or "the frequency (count) of observations in that category divided by the total number of observations" (Agresti & Franklin, 2009, p. 29). Mathematically, a proportion is "two equivalent ratios" (Lobato & Ellis, 2010, p. 7) or "a statement that two pairs of amounts are in the same ratio" (Beckmann, 2018, p. 286). I will use proportion in the statistical sense throughout this study; however proportional reasoning refers to the mathematical definition of proportion.

Reasoning proportionally consolidates many elementary concepts, generally emerges in middle grades, and continues to deepen and expand throughout later years (Lamon, 2012). Proportional reasoning is the comparison of two ratios to determine if they are equivalent or not, and this often involves rates. Following Thompson's (1994) view that a ratio compares quantities in their static states whereas a rate allows application beyond the specific situation, Lobato & Ellis (2010) identified that "a rate is set of infinitely many equivalent ratios" (p. 42). Too often, proportional reasoning is enacted procedurally by both students and teachers (Fisher, 1988; Harel & Behr, 1995;

Izsák & Jacobson, 2017; Orrill & Brown, 2012) by applying the following cross-multiplication algorithm:

$$\frac{a}{b} = \frac{c}{d}, \text{ so } a\ x\ d = b\ x\ c$$

Proportional reasoning is "the ability to recognise, to explain, to think about, to make conjectures about, to graph, to transform, to compare, to make judgements about, to represent, or to symbolize relationships of two simple types … direct … and inverse proportion" (Lamon, 2012, p. 8). Lamon (1993) identified 4 types of proportional reasoning problems: (a) well-chunked, (b) part-part-whole, (c) associated sets, and (d) stretcher and shrinker. Table 1 describes these types and provides examples.

Table 1

Descriptions and Examples for Lamon's Problem Types

Type and Description	Examples
well-chunked problems include a rate that is well known	Speed, price per pound, or incidence/prevalence rates Q: If the incidence rate of the flu is 9%, (there will be 90 new cases of flu per 1000 population) and a town has a population of 100,000, how many new cases of the flu can they expect: *A: 0.09 flu cases per person x 100,000 people = 9,000 cases of flu.*
Part-part-whole problems often give the cardinalities for subsets in terms of a ratio	Ratio of kids to adults or kids' cereal to adults' cereal Q: Kids' cereal and adults' cereal are stocked in a grocery store in the ratio of 5:2. If there are 80 boxes of adults' cereal, how many boxes of kids' cereal is expected? *A: (80 ÷ 2) x 5 = 80 x 200 boxes of kids' cereal.*
Associated sets problems connect two elements that do not have a commonly known relationship	Type of student and a particular response on a survey or handedness and height Q: Of the taller students, 5 are left-handed and 12 are right-handed. Of the shorter students, 5 are left-handed and 18 are right-handed. Are the proportions of left-handed students the same for taller and shorter students?

Type and Description	Examples
	A: No, the proportion of taller students who are left-handed (5/12) is greater than the proportion of shorter students who are left-handed (5/18).
Stretcher and shrinker problems generally have a fixed ratio and measure distance	Changes in height over time or ratio of side lengths for a polygon Q: Consider two radish seedlings that are kept in the same conditions of sunlight and grow at the same rate. At the beginning of day 3 seedling A is 8 mm and seedling B is 10mm. If seedling A measures 14 mm on day 6, what will seedling B measure on day 6? *A: Seedling B will measure 17.5 mm (10 * 14/8)*

Results from this study suggested that students frequently used proportion concepts with associated-sets problems but failed to apply the same concepts to part-part-whole problems and stretcher and shrinker problems. The wording of the questions, not just the type of problem, may have had an impact on students failing to use proportion concepts. These results should be interpreted with respect to the questions that were asked in Lamon's study, which are presented in Figure 4.

These questions and their expected answers are not clearly aligned. Asking which carton has more brown eggs can be interpreted as asking about the number of eggs and not the proportion of eggs. Asking which city has more cars when they all have the same number of cars and different areas and expecting students to consider the density of the cars also seems rather ambiguous. This question can be interpreted as needing to compare the number of cars and would cause students to conclude there is the same number of cars in each city. However, because the question asks which city has more, students may think they have to choose one city over another and then consider comparing cars per square mile.

Figure 4

Problems Used in Lamon's Study

	Part-part-whole
Eggs	The student is shown pictures of two egg cartons, one containing a dozen eggs (8 white eggs and 4 brown eggs) and the other containing 1 1/2 dozen eggs (10 white eggs and 8 brown eggs). Which carton contains more brown eggs?
	Associated sets
Cars	The student is shown two bar graphs, each depicting the number of square miles and the number of cars in two cities. In the first graph, City A has 4000 square miles and 2000 cars, and City B has 5000 square miles and 2000 cars. In the second graph, City A has 4000 square miles and 2000 cars, and City B has 3000 square miles and 2000 cars. In each case, the question is, Which city has more cars?
	Stretchers and shrinkers
Trees	The student is shown a picture of two trees. Tree A is 8 feet high and tree B is 10 feet high. This picture was taken 5 years ago. Today, tree A is 14 feet high and tree B is 16 feet high. Over the last five years, which tree's height has increased most?

Note. From "Ratio and Proportion: Connecting Content and Children's Thinking," by S. Lamon, *Journal for Research in Mathematics Education, 24*(1), p. 44 (https://doi.org/10.2307/749385). Copyright 2017 by the National Council of Teachers of Mathematics.

Similarly, asking which tree heights increased most does not clearly imply how to compare the increases in those trees' heights. Tree A increased 6 feet and Tree B increased 6 feet and because 6 ft = 6 ft, they increased by the same amount, or Tree A started at 8 feet and increased 75% to 14 feet and Tree B started at 10 feet and increased 60% to 16 feet. The question, "Which tree's height increased most?", does not explicitly ask for a comparison of the amount of increase with respect to the initial height. This may explain some of the poor student performance on stretcher and shrinker problems. Questions related to (in)dependence of categorical variables with contingency tables require proportional reasoning and are often unclear.

Proportional reasoning extends beyond comparing the additive difference of two whole numbers. Because ratios are being compared, multiplicative reasoning is necessary. Proportional reasoning includes understanding that a ratio as a composed unit can be iterated or partitioned to create equivalent ratios and that in order to maintain a proportional relationship, if one quantity in a ratio is multiplied or divided by a factor, the other quantities must be treated the same way (Lobato & Ellis, 2010).

Research on Contingency Tables

Research about categorical association dates to 1958 with Inhelder and Piaget, but since then there have been limited studies (Watson & Callingham, 2014). The research I have found with contingency tables includes seminal work in mathematics education with late secondary students in Spain (Batanero et al., 1996), and subsequently a variety of countries (Australia, Germany, New Zealand, Spain, and the United States) and participants (students in elementary, middle, secondary, and undergraduate levels and practicing teachers). The problems with 2x2 contingency tables that Batanero et al. used include the Smoking problem, the Drug problem, and the Allergy problem (see Figure 5). These problems were used as a basis for other studies, sometimes in their original form and sometimes with modifications.

Figure 5

Batanero et al.'s Problems for 2x2 Contingency Tables

A

ITEM 1 (Smoking): In a medical center 250 people have been observed in order to determine whether the habit of smoking has some relationship with bronchial disease. The following results have been obtained:

	Bronchial disease	No bronchial disease	Total
Smoke	90	60	150
Not smoke	60	40	100
Total	150	100	250

B

ITEM 2 (Drug): We are interested in assessing if a certain drug produces digestive troubles in old people. For a sufficient period, 25 old people have been studied, and these results have been obtained:

	Digestive troubles	No digestive troubles	Total
Drug taken	9	8	17
No drug	7	1	8
Total	16	9	25

C

ITEM 3 (Allergy): In order to investigate whether a sedentary lifestyle has some relationship with skin allergy, 30 people have been studied and the following data have been obtained:

	Skin allergy	No skin allergy	Total
Sedentary lifestyle	13	3	16
Nonsedentary lifestyle	2	12	14
Total	15	15	30

Note. Panel A: Smoking problem. Panel B: Drug Problem. Panel C: Allergy problem. From "Intuitive strategies and preconceptions about association in contingency tables," by C. Batanero, A. Estepa, J. D. Godino, and D. R. Green, 1996, *Journal for Research in Mathematics Education, 27*(2), pp. 168-9 (http://doi.org/10.2307/749598). Copyright 2020 by the National Council of Teachers of Mathematics.

Younger participants have different abilities and challenges and whereas older participants are more adept when working with contingency tables, they still struggle with this challenging material. Naïve reasoning with contingency tables might focus on only one of the interior cells, which are the joint frequencies represented by *a*, *b*, *c*, and *d* in Table 2.

Table 2

Contingency Table Structure

		Variable 2		Row Totals
		Column 1	Column 2	
Variable 1	Row 1	a	b	a + b
	Row 2	c	d	c + d
Column Totals		a + c	b + d	T = a + b + c + d

On the other hand, one strategy an expert might use is to look at all four interior cells in a multiplicative fashion to reason about (in)dependence, possibly considering whether the row conditional relative frequencies (RCRFs) are proportional with one another or with the marginal relative frequencies (MRFs). For example, the expert may compare the RCRF, $a / (a + b)$ and $c / (a + c)$ with each other or with the row MRF, $(a + c) / (a + b + c + d)$. For a concrete example with context, consider a variable of age with categories of adult or child and another variable of pet preference with categories of dog and cat (Table 3)

Table 3

Contingency Table for Age and Pet Preference

		Pet preference		Row Totals
		Dog	Cat	
Age	Adult	10	40	50
	Child	20	10	30
Column Totals		30	50	80

Comparing the RCRFs considers the proportion of adults who prefer dogs, which is 10/50 = 1/5, and the proportion of children who prefer dogs, which is 20/30 = 2/3. These proportions might be compared with each other, which leads to an interpretation that adults are less likely than children to prefer dogs as opposed to cats as a pet because the proportion of adults who like dogs is smaller than the proportion of children who like dogs (1/5 < 2/3, comparing the RCRFs). Alternatively, comparing the RCRF with the row MRF considers the proportion of adults who prefer dogs and the proportion of all people (adults and children) who prefer dogs, which is 30/80 = 3/8. This leads to an interpretation that adults are less likely than all people to prefer dogs as opposed to cats as a pet because the proportion of adults who like dogs is smaller than the proportion of all people who like dogs (1/5 < 2/3, comparing the RCRFs). If the RCRFs are equal, the corresponding MRF is the same value and if the RCRFs are unequal, the corresponding MRF is equal to a value between the RCFs.

The research on contingency tables over the past several decades included contingency tables that are complete (with numbers in all cells) and often considered how students conclude whether or not there is (in)dependence. I now trace what we know about reasoning with contingency tables in consideration of age levels and as it relates to my present study.

Elementary Students

Primary school-age children can solve certain contingency table problems; however, they have naive reasoning for all types of tables and tend to either use a limited number of components or use additive reasoning as opposed to multiplicative reasoning (Obersteiner et al., 2015). Inhelder and Piaget (1958) noted younger children tended to use only one or two cells in a contingency table while Shaklee and Mims (1981) found that fourth-grade students predominately used two cells but also used only one cell or additive reasoning with all four interior cells. A recent study revealed that young students often draw the correct conclusions about association or independence but have difficulties justifying their reasoning when ratios and multiplicative rates beyond halving and doubling are required (Obersteiner et al., 2016).

Additionally, in my pilot study, younger students struggled with understanding the structure of a contingency table. For example, they did not see the row marginal frequencies and the column marginal frequencies as representing the same observation. Another study found that younger children performed better on symmetric problems versus asymmetric problems (Saffran et al., 2016). An asymmetric problem structure exists when one category of a variable is the presence of an attribute and the other category of the same variable is the absence of the same attribute. For example, an asymmetric problem might include categories of a variable where fields are treated or not treated with fertilizer. In contrast, if a field is treated with two different types of fertilizer, say fertilizer type A and fertilizer type B, the researchers considered these symmetric problems. In both instances, one category of the variable is the complement to the other concerning the total population of the 2x2 table. These researchers suggested younger students perform better on symmetric problems because the context made the comparison more salient.

Middle Grades Students

Adolescents might possess the ability to reason proportionally but not recognize its applicability to reasoning with contingency tables. Work with adolescents revealed that whereas only 4% of seventh-grade students limited their reasoning with contingency tables to using only one cell when reasoning with contingency tables, the same small percentage were able to use all four cells in a multiplicative manner (Shaklee & Mims, 1981). Most seventh-grade students in this study used two interior cells (25%) or all four cells additively (50%). More recent work (Watson & Callingham, 2015) showed a much larger percentage (compared with 4%) of middle grades students reasoned proportionally on the drug problem where 21% of sixth and seventh-grade students (n = 28) and 30% of eighth and ninth-grade students (n = 21) used all four interior cells multiplicatively. This problem had an inverse association, where not giving a drug was likely to give indigestion, as opposed to a direct association where giving the drug would result in indigestion. Direct and inverse associations are dependent on the table structure and are indicated by the positive or negative value of the difference between the row conditional relative frequencies of the first category of the column variable (Batanero & Sanchez, 2005, p. 272). According to Table 2 (p. 23), assuming giving the drug and having indigestion were in the first row and column respectively, a direct association occurs when the following inequality is met:

$$\frac{a}{a + b} - \frac{c}{c + d} > 0$$

and an inverse association occurs when the following inequality is met:

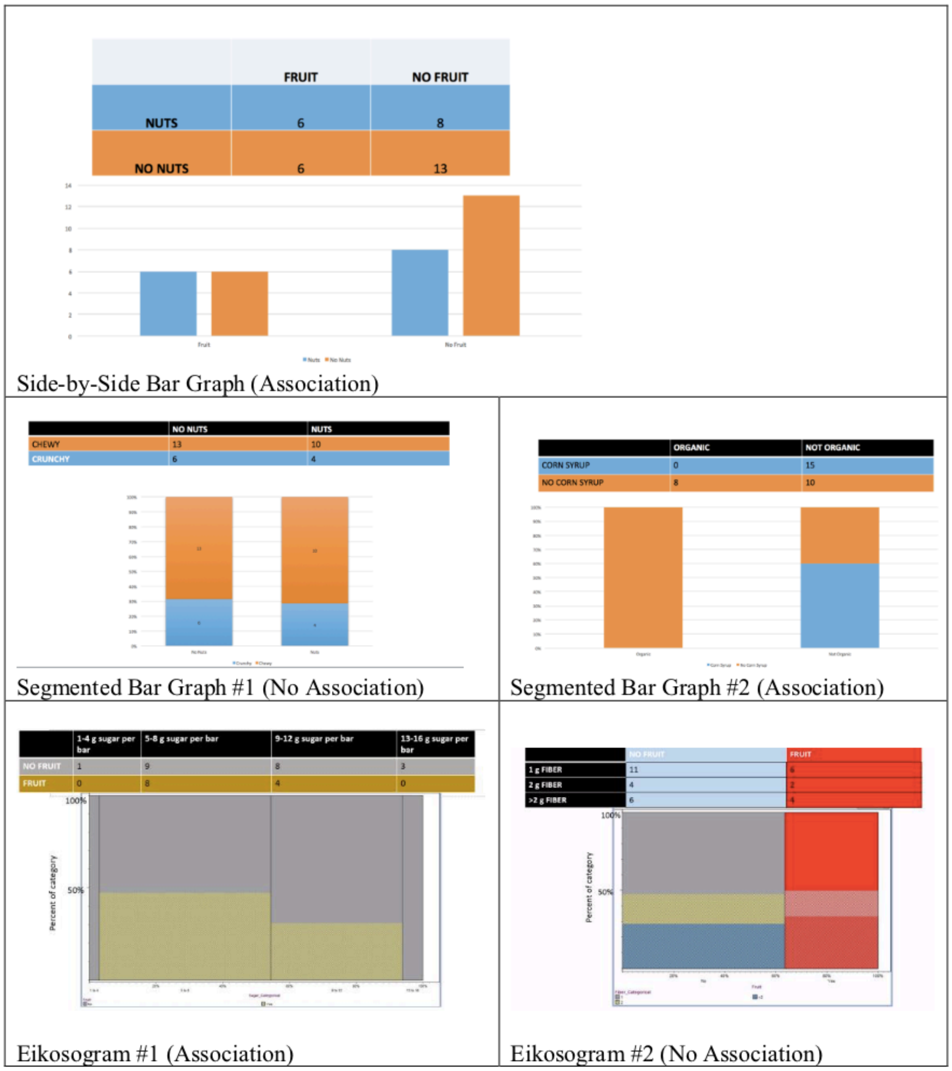
$$\frac{a}{a + b} - \frac{c}{c + d} < 0$$

Most recently, researchers in the United States found that most students aged 11—13 were unable to create a graph showing both variables for categorical data from a contingency table, and those who could make one chose side by side or segmented bar graphs with frequencies rather than relative frequencies (Casey, Hudson, & Ridley, 2018). Results from my pilot study differ in that both a middle and high school student created bar graphs with relative frequencies.

Casey et al.'s study further examined student reasoning about (in)dependence for a data set about granola bars. Five different graphical displays were provided to the students (Figure 6).

Figure 6

Graphs Provided to Students for Granola Bar Problems



Side-by-Side Bar Graph (Association)

Segmented Bar Graph #1 (No Association)

Segmented Bar Graph #2 (Association)

Eikosogram #1 (Association)

Eikosogram #2 (No Association)

Note. From "Students' reasoning about association of categorical variables," by S. Casey, R. Hudson, and L. Ridley, 2018, *10th International Conference on Teaching Statistics*.

These graphs were created from a common data set of granola bars and the participants were shown the data set before reasoning with the graphs. Table 4 summarizes some of the different aspects and the values of attributes for each of the graph problems.

Table 4

Values of Attributes for Granola Bar Problems

	A	B	C	D	E
Graph Type	Side by side Bar Chart	Percent Stacked Bar Chart	Percent Stacked Bar Chart	Mosaic Plot	Mosaic Plot
Dimensions	2x2	2x2	2x2	2x4	3X2
Association	Positive	Independent	Negative	Associated	Independent
Variables	Fruit (Y/N) and Nuts (Y/N)	Chewy/ Crunchy and Nuts (Y/N)	Corn Syrup (Y/N) and Organic (Y/N)	Fruit (Y/N) and sugar per bar (1—4g, 5—8g, 9—12b 13—16g)	Fiber (1g, 2g, >2g) and Fruit (Y/N)
Symmetric Variables	Neither	One	neither	One	One

Casey et al. (2018) concluded the 13 students in the study overall had problems with referent units (attributing the percent to the whole rather than a part) and were hesitant to say there was no relationship. Many students could not use a graph to identify the relationship between the variables. Students who used graphs reasoned about association better with the percentage segmented bar charts than with mosaic plots.

Bar charts are familiar representations and it is likely that middle school students have seen them, although maybe not with a relative frequency scale. Mosaic plots (aka

Mekko charts or Marimekko diagrams) are similar to eikosograms[1] and are based on a unit square where the area is proportional to the number of observations. Some versions of mosaic plots extend the overall area considered to rectangular regions. In either case, considering the area of a square or rectangle connects with geometric concepts that are being developed in middle grades. I believe that these inferior results with the mosaic plot could be remedied by some methodological design improvements. Because the mosaic plot is an unfamiliar representation for most middle school students, some training or time should be provided to allow students to become familiar with its features. diSessa (2004) suggested that students should not be limited to creating the typical sanctioned mathematical representations of tables, charts, and graphs but also be both encouraged and trained in creating other representations and learn profitably from them. The training part may be a key missing factor in Casey et al.'s study, as the students who drew graphs created a traditional graph such as a pie chart, side-by-side bar chart, or segmented bar chart. Additionally, the mosaic plot was the only visual that extended beyond a 2x2 table to include a 4x2 and a 2x3 table, which requires more extensive reasoning. To compare student reasoning between different displays with different dimensions does not seem fair, especially when the one with the simpler dimensions is the more familiar display.

The labeling and numerical differences in the problems are also concerning. The first segmented bar graph is the only one that includes frequency numbers on the graph, which are likely more familiar to students. The difference in the percentages to compare

[1] Oldford (2006) speaks to how eikosograms and mosaic plots are the same in two-dimensions with the exception that eikosograms do not allow for space between areas or "tiles." When the dimensions become larger, the variables alternate axes for mosaic plots; however for an eikosogram the response variable remains alone on the vertical axis and all additional variables are added to the horizontal axis.

(conditional relative frequencies) is much larger for the bar chart with association (60% vs. 0% or 100% vs. 40%) as opposed to the mosaic plot with association (0% vs. 30% vs. 48% vs. 0% or 100% vs. 70% vs. 52% vs. 100%). More concerning is that the width of the mosaic plot for the two categories with no fruit is very small and could easily be overlooked. Another concern is the difference in the labeling of the graphs where the scales on the bar chart facilitated more precise measurement as they included a horizontal line in 10% increments for the vertical axis whereas the mosaic plot only identified 0%, 50% and 100% with only tick marks on the vertical axis. Lastly, the colors in the first three graphs remain the same, whereas the colors for the mosaic plot are different and do not match the contingency table, which is the only legend provided. While the researchers identified some of these concerns, they concluded that students should learn percentage segmented bar charts before mosaic plots and percentage segmented bar charts should be prioritized by software developers. These bar charts are the more traditional graphical displays, but this conclusion seems problematic because this study did not provide a fair comparison of different displays. As researchers suggested, reasoning across these different representations can support some meaningful connections and the deepening of procedural and conceptual understanding (National Council of Teachers of Mathematics, 2014).

Secondary School Students

High school students reason better than younger students but still struggle with employing proportional reasoning to conclude (in)dependence. Seminal work in mathematics education on contingency tables and statistical (in)dependence considered reasoning of late secondary Spanish students (Batanero et al., 1996). This work identified correct, partially correct, and incorrect strategies students used to reason about (in)dependence with a provided contingency table and observed a lack of proportional reasoning. Three incorrect conceptions were identified: (a) *Deterministic conception of association,* where in order to identify association two of the cells (*b* and *c*) must be zero; (b) *Unidirectional conception of association,* where students only identify direct associations ($a \times d > b \times c$) but not inverse associations ($a \times d < b \times c$); and (c) *Localist conception of association,* where students use only a portion of the interior cells in reasoning about (in)dependence. The unidirectional conception occurred when students recognized dependence in the Allergy problem (see Figure 5C, p. 22) but not in the Drug problem (see Figure 5B, p. 22).

The localist conception appeared when a student only considered a portion of the interior cells. For example, in the Smoking problem (see Figure 5A, p. 22), a student claimed there was no dependence because there was a greater percentage of people with bronchial disease who smoke (60%) compared with the percentage of people with bronchial disease who do not smoke (40%). This limited focus on only one of the conditional distributions (e.g. those with bronchial disease) caused students to compare two complementary proportions, which does not address (in)dependence.

In addition to these three incorrect conceptions, the "illusory correlation" was identified through the Smoking problem (see Figure 5A, p. 22). In this problem, the frequencies in the contingency table supported independence of smoking and bronchial disease, but preconceived ideas of a causal relationship where smoking causes bronchial disease prevented the recognition of independence and contributed to claims of association. Subsequent studies with late secondary and undergraduate students also revealed student difficulties in determining (in)dependence with consistent confirmation of the “illusory correlation” (Batanero, Caadas, Daz, & Gea, 2015; Batanero, Cañadas, Estepa, & Arteaga, 2012; Watson & Callingham, 2014). Later studies with university psychology students indicated that student reasoning had only improved for the unidirectional conception of association whereas it has remained the same or regressed for others including the “illusory correlation” (Batanero et al., 2012).

Whereas researchers indicated that first-year university students preferred interpreting numerical as opposed to graphical representations (Batanero et al., 1998; Batanero & Godino, 1998), Glencross questioned this result, noting that representations of data are essential in developing students’ statistical understanding and suggesting students’ limited exposure to graphs had influenced the outcome of the study (Glencross, 1998). He advocated for additional research on this topic.

Context can also influence students’ reasoning with contingency tables in ways beyond the “illusory correlation.” Researchers have found that students’ preconceived theories about the possible causality of the variables can impact their reasoning in different ways (Batanero et al., 1998). Some students only considered the association between the variables if the association could be attributed to a causal relationship between them. For example, in one of the problems where students were ranked by two different judges and there was a moderate correlation, a student claimed there could not be a relation between the order given by the judges because “one judge cannot influence the other” (Batanero et al., 1998, p. 226). “Causal conception of association” (p. 226) includes both when a preconceived causal relationship prevents concluding independence as well as when an expectation of independence prevents seeing an association.

Post-secondary School Students and Adults

Research with Spanish college students showed that overall these students performed worse than secondary students, and most found association when given a scenario similar to the smoking problem that included an expected association based on context but the numbers indicated independence (Batanero et al., 2015, 2012). A few notable exceptions included improvement in identifying inverse relationships and the same level of performance on 2x3 tables in comparison with secondary school students decades earlier (Batanero et al., 1996). Semiotic conflict was cited in both aforementioned studies as well as another study with primary pre-service teachers (Batanero et al., 2015). Godino, Batanero, and Font (2007) defined semiotic conflict as a difference in the meaning of a mathematical expression between the student and the mathematics community or simply as an incorrect interpretation that produces errors. They considered mathematical objects to contain more than concepts and procedures, including language, situations, propositions, and arguments. Semiotic conflict occurs when there is ambiguity in the correct thought, representation, or referent. In the Batanero studies, semiotic conflicts arose when the participants confounded frequencies with relative frequencies, did not separately recognize events and conditions, and did not distinguish between unions and intersections. In one study (Batanero et al., 2015, 2012), two students confounded frequencies with relative frequencies and although they were corrected by the instructors, this more naïve approach of using frequencies persisted for several sessions before being remedied.

Beyond high school, students and adults alike have difficulties with contingency tables and determining (in)dependence. Similar to other age groups, adults were more likely to correctly interpret (in)dependence when the task had a symmetric problem structure and a more grounded context (Osterhaus et al., 2019). A more grounded context includes more familiar and concrete situations as opposed to more abstract or symbolic context. For example, these researchers concluded that college-age students had not fully consolidated the skill of interpreting covariation data because their demonstrated ability was better for a grounded context and symmetric problem structure. With similar results across ages, researchers suggested that the predominance of the more challenging asymmetric variables in past studies may overestimate the difficulty in reasoning with contingency tables (Saffran et al., 2016).

Teachers

Work with teachers and prospective teachers has shown that reasoning with contingency tables is challenging for them as well. Watson and Nathan (2010) found that teachers possess only a partial understanding of the inherent ideas of contingency tables. They suggested that curriculum place more emphasis on categorical data and advocated the importance of helping students understand the association of the variables involved as well as the variables themselves. Many teachers were focused on the mathematics of the contingency table, not recognizing there were statistical variables. More recently, there were some promising results of instructional materials used with teachers and prospective teachers where an item from an assessment of teachers' knowledge for teaching categorical association had a pre-post increase from 69% to 95% (Casey, Ross, Groth, & Zejullahi, 2015).

Mosaic Plots

With the exception of Casey's work, the aforementioned studies generally used only contingency tables as displays. Recently researchers in New Zealand conducted a study using an interactive mosaic plot and found that it may have the potential to assist students to appropriately apply proportional reasoning, especially when considering independence from a probability perspective (Pfannkuch & Budgett, 2017). These researchers conducted an exploratory study and had six first-year college students (age 18—19) who had completed an introductory probability course work with an interactive mosaic plot. All six students were given an individual pre-assessment in which only two students used visual diagrams to solve problems. One of these students and another student had seen a static mosaic plot introduced in the second probability course they were currently taking. The six students worked in pairs with an interactive mosaic plot, neither of the students who had used visual displays to solve problems worked together, and neither of the students who had seen mosaic plots worked together. Thus, two of the three pairs had seen a mosaic plot before. Findings suggested the interactive mosaic plot was useful in helping participants with "proportional reasoning, ability to compare proportions, consideration of proportions in both the horizontal and vertical dimensions; the unlocking and verbalization of simple, conditional and joint probability stories from the data; and visualizing representations for independence" (p. 283). Students attended to features of the display to determine if there was independence or association. Although this research was focused on conditional probability and not statistical association with categorical data and the language between them differs (Watson & Callingham, 2014), it is reasonable to expect the mechanism of the mosaic plot to work similarly.

Summary of Contingency Table Research

People of all ages struggle to use proportional reasoning when working with contingency tables. Younger students can reason correctly with simple contingency tables using benchmark fractions, but they have not developed the necessary proportional reasoning skills to reason coherently across a wide variety of contingency tables. Adolescents and older students may have developed proportional reasoning skills but may not recognize the appropriateness of using relative frequencies (proportions) as opposed to the frequencies (numbers in the contingency tables). Additionally, older students have had more time to develop preconceived ideas based on context, which may interfere with their conclusions. University students and teachers alike struggle in similar ways, and researchers are trying to find instructional strategies and graphical displays that improve student understanding of this important topic.

Theoretical Framework

Because I am interested in considering how students reason about categorical data association, it seems important to address how students learn in more general terms. I believe that a person's knowledge is constructed based on experiences of the individual and their interaction with other people and things. It is not possible to know all the thoughts and understanding that occur within the head of an individual, especially because understanding is "something that is always changing and growing" (Heibert, Carpenter, Fennema, Fuson, Wearne, Murray, Oliver & Human, 1977 p. 4). However, through careful analysis we can gain insights that can be beneficial. Considering knowledge as a synthesis of both empiricism and rationalism, knowledge is formed from experiences and based on reason. How students learn about (in)dependence of categorical variables using contingency tables is influenced by their past experiences, current knowledge, and the information in the contingency tables.

Aspects of Contingency Tables

A contingency table can appear simple on the surface, but it is a representation of complex relationships. Prior research with contingency tables considered many facets that impact reasoning, such as the preconceptions based on the context in the Smoking problem. However, there are additional aspects of contingency tables that have not been evident in past studies. Table 5 summarizes different aspects I considered in this study, the number of studies I found in the literature, and the type of task that researchers suggested are more difficult.

Table 5

Contingency Table Aspects and Findings

Aspect	# Studies	More difficult type
1. Context	5	Contradictory
2. Population Comparison vs. Association of Variables	0	n/a
3. Explanatory & Response Variables	0	n/a
4. Direct and Indirect Association	1	Contradictory
5. Positive and Negative (Inverse) Association	2	Negative (Inverse)
6. Symmetric and Asymmetric	2	Asymmetric
7. Table Dimension & Size	2	Larger
8. Variation	0	Small variation
9. Numbers	0	0

Context

The first aspect I considered in my study and developing tasks was the context defined by the variables and their categories. Context plays a significant role in how students reason about relationships. Because students may struggle with clear meanings for association and independence and may not recognize the statistical variables being compared, I asked questions that are grounded in the context of the categories rather than asking whether they could determine if the two variables are independent or associated. Question posing is different than question asking (Friel et al., 2001), and investigative questions are different than analysis questions. Investigative questions are posed of the entire population and analysis questions are asked of the data (Arnold, 2008). Because contingency tables are created from the data, the questions I asked were analysis questions. Furthermore, these questions were using descriptive rather than inferential statistics. For example, consider a survey question given to middle and high school students that asks them "If you could choose a way to spend the rest of your life, would you choose to be happy or rich?" Rather than asking if they saw association or independence between the variables of student type and life preference, I asked an analysis question like "Are middle school students equally likely, less likely, or more likely than high school students to choose to be happy rather than rich?" As opposed to confounding reasoning with the understanding of more advanced terms, this gave evidence of how participants use their proportional reasoning abilities when working with categorical data in contingency tables.

Participants might have preconceived ideas about some context like smoking and lung cancer that might interfere with their conclusions. When working with completed contingency tables, I did not include problems where there is a conflict between preconceived ideas about (in)dependence due to the context and the determination of (in)dependence using the numbers in the table. The difficulty that secondary and older students have with this contradiction is well documented in the literature (Batanero et al., 2015, 2012, 1996; Casey et al., 2015; Watson & Callingham, 2014) through problems like the smoking problem, and I did not expect younger students to have different results. The relationship between context and prior belief can be classified as theory contradicted, theory supported, and unfamiliar context, and whereas cognitive conflict arises when the theory is contradicted, confidence in an association is increased when the theory is supported (Batanero et al., 1996).

Problems where the theory is supported and association exists may allow students to answer problems correctly without considering the quantities in the table. When considering the context for contingency tables, I aimed to have situations where there is no expectation of (in)dependence and where the situation is of interest to participants (e.g., pet preference for children and adults, type of flu shot given and (in)ability to avoid the flu, type of cereal and shelf location, etc.). Meaningful context is especially important with younger participants, and reducing the interference of expected (in)dependence will allow participants to focus on the quantities and their relationships. For example, a problem could consider the relationship between eating breakfast and test scores, where it is interesting to consider this possible association and plausible that one might exist, but either a determination of independence or association would not be surprising. In this way, the context will motivate but not interfere with the reasoning about (in)dependence.

Whereas the tasks used in my study included contingency tables with situations where there is no expectation of (in)dependence, it may be helpful to have contextual situations where there is a prior theory of (in)dependence for incomplete contingency tables. For example, if a student is asked to complete a table for quantities that have an association and they are not able to do so, changing the context to one that has an expected association (e.g., smoking and lung cancer, exercise frequency and fitness level, etc.) may help advance their reasoning with quantities.

Population Comparison vs. Association of Variables

Another aspect of contingency tables that has been absent in the literature is whether a variable for two different populations is being compared or two variables for one population are being considered. These situations parallel the differences in a chi-squared test for homogeneity or independence, respectively. Information in a contingency table alone does not tell you how the data were collected or how the study was designed. Depending on the sampling scheme, the variables in a contingency table may have one random variable or two. For example, consider a similar survey question of middle school and high school students who were asked, "Given a choice between country music and rap music, what type of music do you prefer?" A statistical investigative question might consider whether one variable is dependent on the other. Three possible sampling schemes could have generated the frequencies in a contingency table. First, both variables could be random variables where possibly I stood in a mall and asked adolescents who passed by whether they were in middle school or high school and whether they preferred country or rap music. In this instance, there are two random variables and a chi-squared test of independence is appropriate. Alternatively, two separate populations determined by either of the categorical variables could have been asked one question (e.g., middle and high school students are asked if they prefer country or rap music; country and rap fans are asked whether they are in middle school or high school). With this situation, there is one random variable and a chi-squared test of homogeneity would be appropriate.

Whereas middle school and high school students need to begin to understand the difference between these designs, this was not a focus of my research. However, it is important to consider this aspect of contingency tables when sequencing tasks from easier to more challenging. I think it is much more unencumbered to reason from a context that includes two different populations that are being compared and furthermore when there is a treatment on one group versus another. This not only explicitly defines the different populations as the categories of one of the variables of the contingency table, but it also creates a clear division between the variables. For example, one variable is the different populations, and the other variable is an attribute of the population such as the answer to the survey question. It seems reasonable to expect that students will be less likely to conflate the quantities for these variables. Thus, when working with participants, I started with problems comparing populations, and as the sequence of tasks progressed in complexity, I used problems that addressed the association of two random variables.

Explanatory and Response Variables

Another factor that I considered when designing tasks with contingency tables is whether there is a clear explanatory and response variable. I have not seen this factor addressed in the literature on statistical association with contingency tables. In the smoking and lung disease problem, smoking is clearly the explanatory variable with a response variable of lung disease. Smoking may explain why a person has lung disease, but lung disease is not the reason someone smokes. Alternatively, a problem considering a possible association of hair color and eye color does not have a clear explanatory and response variable; there is not a causal relationship.

Having explicit response and explanatory variables may have an impact on student reasoning as it does provide a natural way to compare conditional relative frequencies. A natural question is "Compared to non-smokers, are smokers more likely to have lung disease as opposed to not have lung disease?" and to answer this, the different conditions of this explanatory variable are considered separately and compared to each other or the total population. For example, the proportion of smokers that have lung disease might be compared with the proportion of smokers and non-smokers that have lung disease. Without a clear explanatory and response variable, it is reasonable to condition on either of the variables and whichever one you choose, the same conclusion of (in)dependence will result. Because there are more possible relationships of quantities to consider, this type of problem is more complex. Although this is an area that has been absent in the literature, it is beyond the scope of my current research questions. I did, however, consider this aspect when designing tasks for participant interviews and included both types of tasks. I mainly included tasks with explanatory and response variables when evaluating students' ability to understand the structure of contingency tables. These are situations where there is a clear dependent and independent variable rather than where the two variables may be interdependent. I included other tasks where there was not a clear dependent and independent variable toward the end of the interviews, and these tasks allowed me to pose different questions requesting participants to change the variable they condition on.

Direct and Indirect Association

Another contextual consideration is related to the causal relationship. When an association exists, causality is not necessarily present. For example, there might be a spurious correlation of variables where the variable tested is not necessarily what causes the effect. Often times there is a third confounding factor that has a causal relationship. I avoided contexts where there is not a possible causal relationship, as this is not likely to be an engaging and meaningful problem.

When there are causal associations, they can be unilateral (direct) where one variable predicts another or indirect where a third variable may be involved, and this aspect has been considered in past research (Batanero et al., 2015). A direct association exists when one variable directly causes another such as exposure to a disease and acquiring a disease. An indirect association exists when one variable causes a second variable that in turn causes another variable. For example, traveling to a foreign country where a contagious disease is known to be present would have an indirect effect on contracting the disease. In past research (Batanero et al., 2015), the relationship between being an only child and being "problematic" was considered unilateral, meaning the status of an only child directly predicts being "problematic". Additionally, the relationship between a sedentary lifestyle and skin allergy was considered indirect, meaning there was a mediating variable These contexts show how the difference between unilateral and indirect relationships can be subjective, because some people might recognize mediating variables between being an only child and being "problematic," and thus consider it indirect.

Whereas a context with a direct association might be easier to reason with, it may also be more likely to include a preconceived theory. Indirect causal relationships seem like they can vary in how indirect they are, and it is most important to consider that the context is meaningful and an association is plausible. From a context standpoint, I do not think the difference between unilateral and indirect association is an important aspect to consider in this study as it is subsumed by other aspects.

Positive and Negative (Inverse) Association

Whereas past research on how prior conceptions influence conclusions about (in)dependence has been consistent, the same is not true with a positive and negative (inverse) association. Assuming a contingency table with two binary or asymmetric variables, a positive association exists when the presence of one variable is directly associated with the presence of the other variable. (e.g., smoking and lung disease) and the absence of one variable is directly associated with the absence of the other variable (no smoking and no lung disease). A negative (inverse) association exists when the presence of one variable implies an absence of the other variable (e.g., drug is given and no disease is present) and the absence of one variable implies a presence of the other variable. (e.g., no drug is given and disease is present). The direction of the association is dependent on the way the context is presented. For example, if lung health were used instead of lung bronchial disease, what was initially considered a positive association would become a negative association.

I think it is important to include tables with both positive and inverse association, but I do not feel there is a need to have students differentiate between them for the purposes of this study. Although identifying positive and negative association for categorical variables is suggested as an area for future research (Watson & Callingham, 2015), this is largely based on using asymmetric categories for both variables in a contingency table. The results of past studies are mixed where some studies suggested data with a negative or inverse association might be more difficult to identify than data with a positive association (Batanero et al., 1996), but other studies did not detect a difference between identifying a positive or negative association (Batanero et al., 2015).

Whereas differentiating between a positive and negative association for categorical data is not part of the curriculum in the United States, the GAISE Pre-K—12 framework (Franklin et al., 2007) suggested an Agreement-Disagreement Ratio (ADR) for binary categorical data to identify positive and negative association. This is similar to a Quadrant Count Ratio (QCR) that is suggested for quantitative data as a precursor to linear regression. Caution should be exercised when using the ADR. There is an emphasis on quantitative data in statistics and because the calculations for the QCR and ADR are similar, students might think the same approach for quantitative data is always used with categorical data. Once the categorical data is summarized with frequencies (measured quantities) in a contingency table, it is not sufficient to consider a possible linear relationship to determine independence. I used contingency tables with both positive and negative association in this study. Because positive association may be easier, I introduced tables with a positive association between variables before those with a negative association.

Symmetric and Asymmetric

Some past studies have suggested that asymmetric and symmetric context (see Figure 7) can contribute to performance where students have a more difficult time reasoning with an asymmetric context that considers binary variables such as A and not-A (Osterhaus et al., 2019; Saffran et al., 2016).

Figure 7

Contingency Tables with Asymmetric and Symmetric Variables

A

		Plant growing?		Row Totals
		Yes	No	
Fertilizer?	Yes	13	6	19
	No	10	18	28
Column Totals		23	24	47

B

		Plant status		Row Totals
		Growing	Dying	
Fertilizer Type	A	13	6	19
	B	30	18	48
Column Totals		43	24	67

Note: Panel A: Asymmetric variables. Panel B: Symmetric variables.

To uncover students' reasoning about association with contingency tables and not be impeded by types of tables that may be more difficult to master, I started with tasks that had symmetric categories for the variables. Later tasks included one asymmetric variable and two asymmetric variables. Whereas the number and variation of tasks were not structured to provide quantitative analysis or definitive conclusions, it might give some insights as to how these types of tasks influence student reasoning.

Table Dimension and Size

Contingency tables have different dimensions and numbers of categories. A two-dimensional contingency table considers two variables, a three-dimensional contingency table considers three variables, and so on. Past studies that I have seen are limited to two dimensions and are predominated by 2x2 tables–those that have two categories for each of the two variables. Studies that included contingency tables with larger dimensions used 2x3, 3x3 or 2x4 tables. Whereas mosaic plots get more difficult to interpret beyond two variables, they are quite easily extended to more than two categories for each of the two variables. I planned to limit contingency table problems to two-way tables and begin with 2x2 tables. Time did not permit to introduce tables with larger dimensions such as

2x3, and 2x4 tables. Whereas I think there is much to learn about student reasoning with contingency tables with larger dimensions, it is beyond the scope of this study.

Variability

Natural variability, as opposed to measurement or induced variability (Franklin et al., 2007), should be taken into consideration when reasoning with contingency tables. Variability is a key component of statistical reasoning, and a students' understanding of variability develops over time. To date, I have not seen variability addressed in the context of contingency tables.

The information in a single contingency table from a sample can be used to infer (in)dependence for a population. Independence is a property of probability distributions; it is not a sample characteristic. Under independence, the sample proportions have sampling variability and are not necessarily equal. The difference of proportions ranges from 0 to 1, and, whereas neither of these extreme values is expected in practice, a larger absolute value of the difference in proportions indicates a stronger association (Agresti & Franklin, 2009). Additional metrics for the strength of an association include the odds ratio, which uses a part-to-part relationship, and the risk ratio, which uses a part-to-whole relationship.

Some contingency tables cannot have perfectly equal proportions because of the structure of the numbers and the fact that you cannot have fractional observations (see Table 6). In this case, numbers that yield row or marginal relative frequencies that are as close as possible to proportional will signify a very weak association, leading to a conclusion of independence.

Table 6

Contingency Table with Almost Equal Proportions

		Frequent colds?		Row Totals
		Yes	No	
Frequent headaches?	Yes	26	26	52
	No	13	14	27
Column Totals		39	40	79

Additionally, researchers noted that well-developed reasoning acknowledges that some natural variability should be expected (Piaget & Inhelder, 1975; Watson, Callingham, & Kelly, 2007), so when proportions are close to the same there should be a conclusion of independence. There are no apparent guidelines, without conducting a statistical test, to estimate when proportions are close enough to the same to be considered independent. In developing tasks and protocols, I considered this aspect by providing some situations that were close to proportional and asking students how close is close enough.

Numbers and Task Sequence

The numbers used in contingency tables can have an impact on student reasoning beyond the aforementioned aspects. When the joint frequencies are numbers that are easier in calculations such as single digits, multiples of 10, or those that add to a round number like 100, it may unintentionally draw students into additive reasoning. I avoided these situations. I did not want the numbers to be the focus where the context can more easily become detached, so it was important to consider the numbers that were used in the tasks. I began with tables that are easiest to identify independence and association (see Figure 8).

The easiest table to see that independence exists is one where all joint frequency numbers are the same. In this instance, the variables are independent of one another, but the numbers in the interior cells are all equal, so numerically comparing the joint frequencies they may not seem independent. This example allowed students to recognize that they need to reason with measured quantities, not just numbers. Reasoning quantitatively is necessary to recognize that (in)dependence refers to the relationship between the variables and their quantities rather than a relationship among the numbers void of context.

Figure 8

Contingency Tables with Perfect Independence and Association

A

		Life preference		Row Totals
		Happy	Rich	
Grade band	Middle school	37	37	74
	High school	37	37	74
Column Totals		74	74	148

B

		Life preference		Row Totals
		Healthy	Rich	
Grade band	Middle school	74	0	74
	High school	0	74	74
Column Totals		74	74	148

C

		Life preference		Row Totals
		Happy	Healthy	
Grade band	Middle school	0	74	74
	High school	74	0	74
Column Totals		74	74	148

Note. Panel A: Perfect independence. Panel B: Perfect positive association between middle school and a healthy life preference. Panel C: Perfect negative association between middle school and a healthy life preference.

I thought the word association might be difficult for students to understand in a statistical sense. In general, association implies a connectedness or sameness, and with a numerical focus, students might look for numbers to be the same. But in order to recognize statistical association, the variables need to be compared, not just the numbers.

It is easiest to determine that an association exists when a table contains zeros for one of the diagonals. This allowed students who may have a deterministic conception to identify an association, and thus it was a good beginning task for associated variables. I gave this clear example of association twice, alternating which diagonal contains zeros with a perfect positive association between middle school and healthy life preference given first because some findings claim students have an easier time identifying positive rather than negative associations. Even though positive and negative may not have much relevance because the variables may not be asymmetric, students may be likely to look for the largest number in the first cell, so keeping this order was reasonable. Throughout the tasks, I asked probing questions to determine what cells and relationships students are using to decide (in)dependence.

The next group of problems included problem pairs where first the row marginal frequencies are equal and second neither row nor column marginal frequencies are equal. Past findings indicated that some students think the marginal frequencies need to be equal to claim independence, so this type of variation in problems is warranted. The first pair in this group focused on independent situations where the first problem has larger numbers in the second column and the companion problem had larger numbers in the first column (See Figure 9A). I expected the companion problem to be more difficult because students might compare the number of observations rather than the proportion, which would result in an incorrect conclusion and indicate the student was using additive rather than multiplicative reasoning.

The second pair of problems in this group (see Figure 9B) focused on situations with variables that have a positive association with categories in the first row and column. Both problems had the largest numbers in the interior cells of the contingency table contained in the first diagonal (cells *a* and *c*). These problems are ones where additive reasoning using all four cells leads to a correct conclusion, and because my work with participants was more than simply an observational interview, I probed participants to explain their reasoning to see if they recognize the multiplicative relationship.

Figure 9

Companion Problems with Same and Different Marginal Frequencies

A

		Music preference		**Row Totals**
		Rap	Rock	
Grade band	Middle school	27	47	74
	High school	27	47	74
Column Totals		54	94	148

		Music preference		**Row Totals**
		Country	Pop	
Grade band	Middle school	75	19	94
	High school	43	11	54
Column Totals		118	30	148

B

		Music preference		**Row Totals**
		Country	Rock	
Grade band	Middle school	42	32	74
	High school	27	47	74
Column Totals		69	79	148

		Music preference		**Row Totals**
		Rap	Country	
Grade band	Middle school	49	32	81
	High school	23	44	67
Column Totals		72	76	148

Notes. Panel A: Independent problems. Panel B: Problems with a positive association between middle school and a rap music preference.

When the largest numbers in the interior cells of the contingency table are not contained in one of the diagonals, additive reasoning cannot be used to support a correct conclusion. The remaining problems did not include larger numbers on a diagonal, and therefore I expected them to be more challenging for participants. A mix of (in)dependent problems were given beginning with equal marginal frequencies and increasing in difficulty.

Analytical Framework

My study included think-aloud clinical interviews of students working with tasks I designed based on the literature and the different aspects I discussed previously. The data corpus consisted of video and audio recordings, transcripts, lesson graphs, scanned images of student work, and field notes. To analyze the interviews, I developed a framework based on past studies and a pilot study I previously conducted. The framework includes nine conceptions of reasoning with contingency tables (see Table 7), which are based on the five levels (L1—L5) identified by Perez-Echevarria (1990, as cited in Batanero et al., 1996). These levels are not necessarily developmental levels but are distinct types of reasoning that are likely influenced by students' proportional reasoning abilities. Levels L1—L4 from Perez-Echevarria directly align with my L1—L3 and A1, where I named the first 3 localist (L) conceptions and the last one additive (A). Additionally, like Watson and Callingham (2014), I included a category to account for students who used no cells of the contingency table for their reasoning (N0 and N1). Although these researchers separated this category into four parts (0-no justification, 1-idiosyncratic, 2-personal opinion, 3-survey info) I choose to use one category (N0). Because the task design attended to an aspect of context, I did not expect to see any N0 reasoning. On the other hand, I separated L5 into three different categories (P1—P3) based on Batanero et al.'s (1996) specification of different theorems in action.

Table 7

Initial Framework for Reasoning with Contingency Tables

Code	Name	Description and features
N0	No interior cells used and 0 marginal values used	No cells in the table are used. Students may use their preconceived notions or other reasoning about the context to decide about (in)dependence.
N1	No interior cells used, but one or more marginal values used	Only the exterior cells are used. Students may use the structure of the table. Students may realize marginal frequencies are limits.
L1	Localist, 1 interior cell used	Only one cell in the table is used to decide about independence or association. This is likely to happen when the student focuses on the largest value in the table, especially if it is in the first cell.
L2	Localist, 2 interior cells used	Two cells in the table are used to decide about independence or association. This typically includes comparing *a* vs. *b* or *a* vs. *d*.
L3	Localist, 3 interior cells used	Three cells in the table are used to decide about independence or association.
A1	Additive, 4 interior cells used	All four cells in the table are used to decide about independence or association, but it is only done in an additive way (e.g., because *a* and *d* are bigger than *b* and *c*). This includes deterministic reasoning where *b* and *c* are thought to be restricted to 0 for an association to occur.
P1	Proportional, risk ratio reasoning with interior cells	All four cells in the table are used to decide about independence or association and multiplicative reasoning is used. Proportional reasoning compares risk (part to whole ratios) and **compares one conditional relative frequency to another** focusing on the interior cells.

Code	Name	Description and features
P2	Proportional, risk ratio reasoning with interior and exterior cells	All four cells in the table are used to decide about independence or association and multiplicative reasoning is used. Proportional reasoning compares risk (part to whole ratios) and **compares one conditional relative frequency to a marginal relative frequency**, using both interior and exterior cells.
P3	Proportional, odds ratio reasoning	All four cells in the table are used to decide about independence or association and multiplicative reasoning is used. Proportional reasoning compares odds (part to part ratios) and **compares the odds for one category to another category for the same variable** through subtraction or a ratio, focusing on the interior cells.

When considering the problems where the mosaic plots were provided, I used the same codes and appended a code to address how the mosaic plot seemed to function. In my pilot study, I considered whether mosaic plots were a hindrance (M-), seemed to have no impact on a solution (M), or were helpful (M+). The mosaic plot was never seen to be a hindrance, possibly because the participants were instructed to create one in the pilot study. I expected the depth of understanding and ways in which participants used the mosaic plot to vary, and it was useful to further refine the M and M+ codes. Friel, Curcio, and Bright (2001) noted three levels of graph comprehension – elementary, intermediate, and advanced. A student with elementary comprehension reads the data and identifies information from the graph in context, and considering a mosaic plot this could be evidenced by understanding the total area represents all observations, one of the bars represents a condition of one of the variables and one of the tiles represents a joint frequency. A student with intermediate comprehension reads between the data and finds contextual relationships. For a mosaic plot, this might be revealed by comparing the sizes and relative sizes of the tiles that represent the joint frequencies. A student with advanced comprehension reads beyond the data, demonstrating an ability to succinctly summarize the graph and use it to answer questions that may extend beyond the information provided. Using a mosaic plot to reason about association demonstrates advanced comprehension. Accordingly, I will use M1, M2, and M3, which align with these levels and are summarized regarding the use of a mosaic plot in Table 8.

Table 8

Initial Framework for Reasoning with a Mosaic Plot

Code	Level	Use of Mosaic Plot
M1	Elementary	Identifies individual parts related to context
M2	Intermediate	Recognizes relationships of parts in context
M3	Advanced	Reasons in context about (in)dependence

This framework was a place to start analyzing the data, but it was not simply something that was applied to the data. The framework and the data were both continuously interrogated with one another, and I made modifications to the framework throughout the analysis. This initial framework changed; thus the resulting framework is something that emerged from the data, so it did not just assimilate the data, but rather accommodated the data.

CHAPTER 3

METHODOLOGY

To better understand how students reason about (in)dependence of categorical variables when using contingency tables and mosaic plots, I conducted a series of clinical interviews with middle and high school students. I developed a protocol for each interview to guide the questioning. I based these protocols on past research and each interview informed the subsequent protocols. I analyzed the information resulting from the interviews, including video recordings, transcripts, and written work in consideration of the research questions:

1. In what ways do students reason about (in)dependence of categorical variables when using contingency tables?
2. In what ways do students use mosaic plots to reason about (in)dependence of categorical variables when using contingency tables?

I started with an initial framework to analyze the data. The analysis process included making modifications to the framework in order to account for emergent themes that could not be captured by current forms of the framework. The result of which is a revised framework I present in this dissertation.

Because I was interested in understanding student reasoning rather than evaluating a final answer in written form and my research questions were of the "how?" and "why?" nature, a qualitative study was appropriate (Denzin & Lincoln, 2005). I used a case study methodology, which is commonly used in education research (Yazan, 2015), and combined case study perspectives of prominent methodologists. Whereas many social scientists consider case studies to be valid only for exploration, Yin (2018) argued they can be valid for explanation as well as uncovering possible causal relationships. Yin (2018) identified case study methodology as being advantageous when these "how" and "why" questions are related to "a contemporary set of events over which a researcher has little or no control" (p. 13). The context for this study represents a contemporary set of events because scant research exists about how younger students reason about (in)dependence with contingency tables and mosaic plots and most school standards include reasoning with bivariate data using multiple representations. As a researcher, I have little control over what students have encountered in their mathematics instruction or how they are reasoning. As suggested by Stake (1995), I started this case study with two sharpened research questions that helped to structure the interviews and I used a review of relevant literature to construct my theoretical framework (Merriam, 1998).

I selected a multiple-case embedded design (Yin, 2018). According to Merriam (1998), a case is a single thing, and for my study, a case is a participant. Using multiple cases allows me to analyze the data both within and across participants (Yin, 2018). My design is embedded (Yin, 2018) because there are multiple interviews for each participant and multiple tasks within each interview. Additionally, I not only considered the participants' reasoning with complete and incomplete contingency tables, but I also considered their reasoning with mosaic plots.

As a researcher, throughout the interviews, I was placed in social interaction with participants and this allowed me to "experience" the students' mathematics through experiencing constraints in interacting with them. A qualitative methodology allowed for an in-depth, detailed study of issues that was not accessible through quantitative studies alone, and these qualitative methods have become "increasingly important modes of inquiry for social sciences and applied fields" (Marshall & Rossman, 2016, p. 1). Qualitative researchers analyze details and build an intricate and holistic picture (Creswell, 2009), and they have a "process orientation toward the world" (Maxwell, 2013, p. 30). Through a qualitative approach, I was able to see more fine-grained elements than if I were to survey participants through a questionnaire or consider students' performance on written work whether it is a standardized test or some other form of written assessment. Patton (2002) stated that qualitative researchers "provide a framework within which people can respond in a way that represents accurately and thoroughly their points of view" (p. 21). This detailed information can aid in understanding the complexities of problem solving and cannot be obtained through surveys.

Interviewing

Seidman (2013) suggested that if a researcher is interested in a student's experience and the meaning they attribute to that experience, then interviewing is the most appropriate method of obtaining data. As opposed to a questionnaire or test that can more easily be given at scale and the accuracy of answers can be considered quantitatively, an interview can help the researcher reveal a more fine-grained understanding of student reasoning and uncover the motivations for the answers. The strengths of more open-ended techniques such as clinical interviews include "the ability to collect and analyze data on mental processes at the level of a subject's authentic ideas and meanings, and to expose hidden structures and processes in the subject's thinking that could not be detected by less open-ended techniques."(Clement, 2000, p. 341).

Clinical Interviews

To understand how students reason about (in)dependence with categorical data, I conducted clinical interviews, which are an effort to uncover the participants' natural thoughts (Clement, 2000) and witness their understanding knowledge, thoughts, and development at a finer level of detail than what structured interviews, observation, or testing alone can reveal (Ginsburg, 1997). I used a structured interview, which had a set of predetermined and standard questions that were the same for each participant only to initially evaluate prerequisite skills. All subsequent interviews were semi-structured, where there were defined topics and some planned questions; however, I asked other questions dependent on the actions of the participant. Ginsburg (1997) suggested that clinical interviews are "deliberately nonstandardized" (p. 2) and are an effective way to obtain "a rich and sensitive view of cognitive processes" (p. 26). All the interviews were clinical interviews where I treated each participant differently according to their actions and my perception and understanding.

I began each of the interviews with an open-ended task where the aim was to allow the participant to speak for themselves and "structure the task in any way he sees fit" (Ginsburg, 1981, p. 6). Some researchers define an open task to include both open-ended and open-middle tasks (Yeo, 2017). An open-ended task has multiple correct answers, whereas a closed task has one solution, and an open-middle task has one correct answer but multiple solution methods. Open-middle tasks, like open-ended tasks, cannot be solved with a routine procedure and may require a considerable amount of checking to verify sense-making and question-formulating when solving (Bell & Burkhardt, 2002). I used primarily tasks that are open-ended and open-middle as both types of tasks allow the participant to choose how to approach and solve the task.

Based on Piaget's work, Ginsburg (1981) identified three aims of exploring the mathematical mind of children through clinical interviews: (1) discovery of the thought processes, (2) identification of thought processes, and (3) evaluation of competence levels. The goal of the interview is not to define thought processes a priori but rather to "observe, explore and attempt to discover" (p. 5) thought processes. The technique for the first aim of discovering the thought process is to begin with an open-ended task. Additionally, an interviewer should focus on revealing rather than leading the participant's nascent thoughts by following-up in a contingent manner and requesting reflection on both how (process) the answer was determined and why (rationale) a particular method was used. Secondly, when identifying thought processes, the interviewer should focus on the "interesting intellectual phenomena" (p. 6) that have been discovered, again using open-ended tasks, and contingent follow-up and reflection. These phenomena could conceivably come from literature or previous interviews. When considering a phenomenon, the researcher should identify all possible solution methods and begin the clinical interview by directing the participant's behavior to the phenomena, often with an open-ended task. The clinical interview aims to (a) "facilitate rich verbalization" (p. 7), (b) verify and clarify statements, and (c) test "alternative hypotheses concerning underlying processes" (p. 7). This testing may come through contingent questioning, which is purposeful but not necessarily standardized. Finally, the evaluation of competence levels is not determined by simply the best reasoning a student exhibits one time, but rather the student's ability to reason flexibly across a variety of structures and contexts in a consistent manner. Clinical interviews are a crucial tool for mathematics education researchers because they allow insight into students' problem-solving strategies and reasoning (Schoenfeld, 2002).

Think-aloud Methods

To achieve these goals of understanding student reasoning, I asked participants to think-aloud while they engaged in tasks. Think-aloud methods promote that participants explain their reasoning while working on a task and have a "sound theoretical basis and provide a valid source of data about participant thinking"(Charters, 2003, p. 68). Whereas some researchers model a think-aloud activity for instructional purposes, Gibson (1997) noted doing so might introduce bias by inducing the participant to model instructions rather than elicit their spontaneous thoughts, so I did not do this. Using a think-aloud method allowed me to engage in a type of assessment by identifying a participant's thinking at a particular point in time.

I used these think-aloud methods in interviews where my role ranged from observational to probing to assistive. I developed these stances based on variations in observer involvement (Patton, 2002). Observational interviews allowed the participants to work through the problems and I mostly asked the participant "What are you thinking" to remind them to think-aloud. I responded to their questions by restating their question or asking them what they think the answer may be. Alternatively, during a probing interview, I posed clarifying questions when there was something that was unclear or may benefit from a more in-depth explanation. Additionally, I used questions to gain insights into the participant's confidence level and range of thinking including alternative ways to solve the problem. Finally, during an assistive interview, which aims to advance student reasoning, I provided instruction or explanation about some new material and posed questions that lead the participant to consider certain information or a different approach.

Another way in which think-aloud interview techniques vary is the timing of the thinking aloud. As the interviewer, I primarily asked the participants to think-aloud as they were solving the tasks, which is a concurrent approach. Alternatively, with a retrospective approach, the interviewer waits until the student has finished solving the task and then asks them to explain their thinking. Kuusela & Pallab (2000) compared a concurrent and retrospective think-aloud protocol, and their findings favored the concurrent protocol because it resulted in more and richer observations; however, the retrospective protocol had advantages of better data about the final decision made by the participant. Because the retrospective approach allows the final answer to taint the recall of the thought process and I was more interested in the participant's moment-to-moment reasoning, I mainly used a concurrent think-aloud approach for all interviews. However, there were times when I felt that interrupting a participant's work to ask what they were thinking might derail the strategy that was developing, so, at times I used a retrospective approach.

Summary of Interviews

I created a series of 5 clinical, think-aloud interviews that begin with verifying the participants' prerequisite ability to reason with contingency tables and progressed to consider participants' work with contingency tables (see Table 9). The remaining interviews revealed participants' current understanding of completed contingency tables, introduced multiple representations, and ultimately included work with incomplete contingency tables.

Table 9

Interview Attributes

Number	Goals	Interviewer Role	Interview Type
IV#1	Verify necessary reasoning (proportional, probability, CT structure)	Observational	Structured
IV#2	Reveal contingency table reasoning	Probing	Semi-Structured
IV#3	Assessment of Venn diagrams; introduction to and reasoning with mosaic plots	Observational, Probing & Assistive	Semi-Structured
IV#4	Reasoning with contingency tables with missing values	Assistive	Semi-Structured
IV#5	Reasoning with contingency tables with missing values and mosaic plots	Assistive	Semi-Structured

Study Design

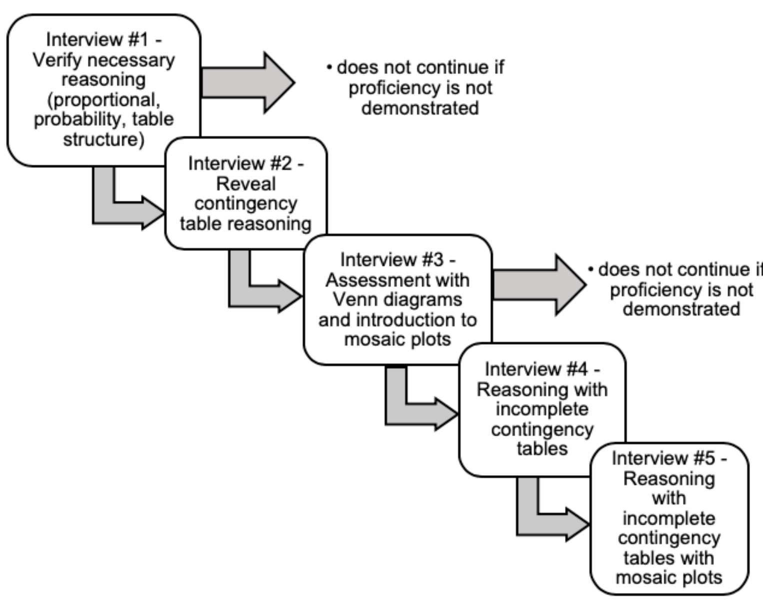
Recalling that my research questions aim to uncover how students reason about (in)dependence of categorical variables in the context of contingency tables with and without mosaic plots, I conducted a series of clinical interviews. I carefully selected the participants based on relevant criteria and used tasks that were designed to elicit and advance reasoning. The criteria for participant selection was informed by a pilot study I completed.

Because I wanted to get a sense of ways that students across upper elementary, middle, and high school reasoned with contingency tables and mosaic plots, my pilot study included one interview with seven participants who ranged in age from 7 to 17 years old. The younger students did not have the proportional reasoning or understanding of the structure of a contingency table that is necessary to reason across a range of problems. Students overall were able to construct a mosaic plot and showed improved reasoning when a mosaic plot was provided, but they still struggled on some of the problems, especially when the contingency tables were incomplete. The words association and independence and coordinating constituent components created challenges, and students did not recognize the applicability of the distributive property. I used the findings of my pilot study to inform the development of tasks and interview protocols and to narrow the grade range for my participant pool.

These encounters with participants were the basis of the data corpus and included a series of interviews that ranged from observational to assistive. Some of the interviews (IV#1 and IV#3) required participants to demonstrate proficiency to continue with subsequent interviews (See Figure 10).

Figure 10

Interview Sequence for Participants



Participant Selection

I began seeking eight initial participants and planned to conduct a series of up to five interviews with each participant for a maximum number of 40 interviews. I chose initial participants from my community, privileging those I already knew through my neighborhood, church, or other community connections and those I judged to be metacognitively aware based on our past interactions. I aimed to find students who already understood the structure of a contingency table and had an understanding of proportional reasoning where they could apply it in context. I wanted to recruit participants who may have been exposed to contingency tables but have not extensively studied contingency tables and statistical methods of determining associations such as chi-squared or Fisher's exact tests. Thus, I did not consider any students who have taken AP Statistics because their reasoning about contingency tables and (in)dependence is likely to be more developed.

Because elementary students are not likely to reason proportionally, I considered students in seventh grade and above who have completed courses addressing proportional reasoning, which are typically taught in sixth and seventh grade according to the approved curriculum in Georgia. Students in these grades are more likely to have the desired proportional reasoning abilities as compared with younger students who may not have been exposed to this material in an educational setting. In a recent study (Riehl & Steinthorsdottir, 2014) middle grades students were progressively less likely to incorrectly solve a missing value proportion problem with illogical errors holding at about 22% and additive errors decreasing from 60% in grade 6 to 33% in grade 8. Whereas this may indicate that over half (22% + 33%) of eighth-grade students are not able to solve this type of problem, this was a response to a paper and pencil test and not necessarily indicative of proportional reasoning abilities necessary to determine (in)dependence with contingency tables. Nonetheless, not all students in seventh grade and above can reason proportionally, so I used the first interview to screen for this ability.

I requested participation from the parents of 14 students of whom 6 students declined to participate, primarily due to busy schedules and interviews being conducted outside of school hours. All 6 of these non-participants were girls, and they were in seventh, eighth, ninth and 11th grades. The resulting eight participants (P1—P8) were students from a suburban public middle or high school, and three of the participants (P3, P4, and P5) were siblings. Five of the eight participants were students who reasoned proportionally and completed all interviews (S1—S5). Table 10 summarizes the participants.

Table 10

Summary of Interview Participants

Participant/ Student Number	Pseudonym	Grade	Gender	Interviews Completed	Reason Proportionally
P1/S1	Jamie	8	M	8	Y
P2	Jordan	11	F	2	Y
P3/S2	Zander	8	M	8	Y
P4/S3	Sydney	9	M	8	Y
P5	Harper	8	F	8	N
P6/S4	Hayden	9	M	8	Y
P7	Cameron	8	F	1	Y
P8/S5	Jessie	7	M	8	Y

Data Collection

There were up to five interviews for each participant, and each interview lasted for approximately one hour. This gave sufficient time to uncover students' reasoning but for the most part, the interviews were not so long as to allow mental fatigue to interfere with their reasoning. To improve reliability, I video-recorded all interviews with two cameras, capturing separate views as suggested by Hall (2000). One camera captured the written work of the participant, and the other camera had a wider-angle view to include a full view of the participant. Videotaping from two perspectives allowed a narrow perspective where much has been deleted but specificity and clarity emerge. In a wider view, what was deleted in the narrow view has been restored, so eye movements, facial expressions, gestures, and other influences in the larger environment can be considered. Video recordings have limitations. For example, they may cause the participants to be less comfortable, but because my cameras were static with no additional person operating them, I feel this effect was limited, although it may have been the reason that some people declined to participate or dropped out of the study. Video recordings have limitations in that not everything can be seen because of the angle of recording. Using two views mitigated this and the combined views provided a fine-grained, multimodal, sequential record that is durable and sharable (Jewitt, 2012).

I recorded comments on the written protocol during the interviews and completed field notes directly after the interview as well as throughout the study. I aimed to schedule the interviews with each participant approximately a week apart to allow time for transcription and analysis before the next interview. Due to scheduling challenges, the time between interviews was sometimes shorter and sometimes longer than a week.

After video recording each interview, I retained the written student work and scanned it to a pdf file for accessibility (see Appendix A for Data Management Procedures). An audio file was created from the video recording using iMovie (Version

10.1.14), and an automated speech to text transcription application, Otter (Otter.ai) was used to generate a transcript that includes speaker identification and timestamps. I used Otter's online tools to edit the transcript for accuracy and exported this to a text file. Inqscribe (Version 2.2.4.262) was used to transform time stamps into a format compatible with ATLAS.ti (Version 8.4.4), and a Word file was created for each transcript. The two videos were combined for a single view (picture in a picture) with both near and far images. The combined view video and transcript were imported into ATLAS.ti and linked.

I reviewed each combined video in conjunction with the transcript. This resulted in an annotated transcript that identifies participants' meaningful gestures and actions, including what they are writing or drawing. Additionally, I created a lesson graph, (Seago, 2003) of the data, simultaneously working with the video recordings, transcripts, and scanned work to summarize the events of the interview and my observations in contiguity. The data corpus in its entirety is summarized in Table 11. The students' scanned work was loaded into ATLAS.ti, which was used for analysis.

Table 11

Data Corpus Number of Items

Data Type	IV#1	IV#2	IV#3	IV#4	IV#5	Total
Video 1 (close-up)	8	7	6	6	6	27
Video 2 (wide-angle)	8	7	6	6	6	27
Audio Files	8	7	6	6	6	27
Transcripts	8	7	6	6	6	27
Combined video	8	7	6	6	6	27
Scanned work	8	7	6	6	6	27
Annotated Transcript	8	7	6	6	6	27
Lesson Graphs	8	7	6	6	6	27

Interviews

My interview protocols were based on Charters' (2003) suggestion of using intermediate-level tasks. If a task is too easy, the participant will solve it without struggling, and their reasoning will be more abbreviated, not giving much insight. If the task is too hard, the participant may not be able to reason about it at all, which provides no insight, or they may become too frustrated and want to stop. After the initial interview to assess the necessary skills of proportional reasoning, probability, and contingency table structure, the following interviews all focused on reasoning about association of categorical variables using two-way contingency tables. Throughout the problems in all interviews, I did not use the words association or independence because I was concerned that these younger students might have difficulty with them. Instead, I used the words less likely, equally likely, and more likely or asked questions about association in the context of the problem. I did not use the words category or variable for the same reason; instead I used the context of the problem. All five interviews are summarized in Table 10 on p.66.

Interview 1 (IV#1). Because proportional reasoning is needed to coherently determine (in)dependence with contingency tables (Inhelder & Piaget, 1955, as cited in Batanero, Estepa, Godino, & Green, 1996), I used the first interview to assess students' proportional reasoning. The goal for this interview along with the tasks and answers are included in Appendix B. At the time of the first interview, I was not interested in how their reasoning was changing or how their knowledge was advancing but rather their current abilities. Thus, this first interview (IV#1) used an observational think-aloud approach, primarily asking "What are you thinking?" or requesting clarifications of their statements. In addition to evaluating their proportional reasoning, I evaluated their ability to understand simple probability tasks and the basic structure of a contingency table. One participant, Harper (P5) did not demonstrate the desired understandings of proportional reasoning and probability and was not considered or analyzed for the purpose of this study.

I used proportional reasoning tasks to assess their ability to determine if two ratios of measured quantities are equivalent. This was assessed by using missing value and multiplicative comparison problems in context. Because I was not interested in assessing numerical fluency and computational skills, a calculator was available to all participants throughout the interviews. Additionally, I used probability tasks to assess if students could clearly define the event and sample space as well as redefine the sample space to calculate a direct probability and reason with it in context. Finally, I assessed their understanding of a contingency table structure, which includes recognizing this structure both quantitatively and numerically. Quantitatively, this includes recognizing that each joint frequency is a summary statistic that counts observations with a category of one variable while simultaneously counting observations with a category of the other variable. Row and column marginal frequencies are two different ways of categorizing the same total number of observations where row marginal frequencies include both categories of the column variable and column marginal frequencies include both categories of the row variable. Numerically, this includes recognizing that joint frequencies are mutually exclusive and jointly exhaustive, marginal frequencies are partitioned into joint frequencies, and the total frequency is partitioned into two sets of marginal frequencies and one set of joint frequencies. While the values in contingency tables add across and down, the total is also the sum of the interior cells. Thus, I gave students contingency tables with missing values that can be determined through row or column addition. I then asked if the same observation can be included in the numbers for both a row marginal frequency and a column marginal frequency.

Interview 2 (IV#2). The first of the interviews focused solely on contingency tables (IV#2) used an observational think-aloud approach to reveal the students' current, unaided reasoning about (in)dependence. The goal for this interview along with the tasks and answers are included in Appendix C. The tasks in IV#2 progressed in their difficulty and questions used the context of the categories of the variables. This interview uncovered whether they saw proportional reasoning as applicable in determining (in)dependence and if there was a difference in their reasoning when given different scenarios. Due to time limitations, not all aspects of contingency tables were separately considered. This interview and subsequent interviews (IV#3—5) focused on reasoning with contingency tables and included tasks designed around the research questions and in consideration of the literature. The aspects of contingency tables were used to design a sequence of real or realistic problems that progress in difficulty.

Interview 3 (IV#3). The next interview (IV#3) introduced alternative visual displays and began with a problem that used the same context as a problem in IV#2. The goal for this interview along with the tasks and answers are included in Appendix D. I first requested the participant to complete a Venn diagram from the information in the table. I provided a skeletal Venn diagram and expected students to be agile with these representations because they are widely used across subjects in school. The participants struggled to label the parts and place the numbers, so I provided a Venn diagram with the numbers from the four interior cells of the contingency table included on the diagram.

Although students were able to label this correctly, only one student continued to use this approach when asked to draw another Venn diagram for a different problem. Whereas these unexpected challenges with Venn diagrams reveal interesting phenomena, it is beyond the scope of this study, and these portions of the interviews were not analyzed.

I next instructed the participants to complete a mosaic plot, because it was not likely that the participants had seen one before. Researchers encourage the use of multiple representations noting "the ability to deal with them flexibly is key to successful mathematical thinking and problem solving" (Dreher & Kuntze, 2015, p. 91), however, they also caution that students need to be "encouraged to actively create connections between these representations" (p. 91). I guided them through the process of creating a mosaic plot, providing written instructions, an example of a computer-generated mosaic plot, and an example of a mosaic plot drawn on a 10x10 grid. I provided them a blank 10x10 grid and used both numbers and context while being sure to include appropriate labels. I first had them draw a corresponding mosaic plot for a contingency table that represents an association of variables. Next, I requested participants to create a mosaic plot for a second problem that has independence, and I provided assistance where needed. This gave them experience creating a mosaic plot for both types of (in)dependence. Next, a series of problems with completed contingency tables and accompanying mosaic plots were provided. These problems mirrored the problems in IV#2, although some participants completed more or fewer problems in this session, so there was not necessarily a companion problem across these interviews. This interview ranged from more observational to probing to assistive.

Interview 4 (IV#4). For the next interview (IV#4), problems included contingency tables with missing values that required proportional reasoning to complete correctly. The goal for this interview along with the tasks and answers are included in Appendix E. Participants were given conditions of (in)dependence, and accompanying mosaic plots were not provided. Problems with incomplete contingency tables have been absent from the literature as far as I have seen. This interview began with problems where all values were missing except for the total number of observations. Next, problems and became progressively more challenging as values for marginal frequencies and joint frequencies were included. Aiming to learn mathematics more deeply, researchers encourage reversibility questions, which are characterized by giving the answer and asking for the question (Dougherty et al., 2016). To fully understand how students reason with categorical data in contingency tables, it is important to consider how they might choose numbers to make the situation (in)dependent.

Interview 5 (IV#5). The final interview (IV#5) first allowed the participants to draw a representation of their choosing for a provided contingency table. The goal for this interview along with the tasks and answers are included in Appendix F. Next, I requested them to draw a mosaic plot if that was not the representation they chose. Then, I asked the students about their understanding of some words (association and independence; category and variable) and how they saw them in the contingency table and mosaic plot. The remaining problems were the same or similar to the previous interview; however, a mosaic plot accompanied the contingency table.

Data Analysis

While analyzing the data, I started with the aforementioned framework as a basis for coding interview transcripts and videos. My goal was to analyze the data to understand how students reasoned in order to answer both research questions. I included the 5 participants who completed all the interviews and demonstrated proportional reasoning in the first interview. Interviews #2 – 5 were included in my analysis because they were aimed at assessing reasoning with contingency tables and mosaic plots. Thus, I analyzed a total of 20 interviews.

I began by analyzing the second interview which aimed to answer the first research question and included tasks that are a basis of the subsequent interviews. I reviewed the proposed framework, and with it in mind, I examined the videos, transcripts, and student work for each participant. For each task I considered the possible available codes and the student reasoning. There were not any instances of reasoning with no cells or additive approaches (N0, A1). There were a limited number of instances of reasoning with limited interior cells (N1, L2) or using a mix of conditional relative frequencies and marginal relative frequencies (P2). This is likely because of the way the tasks were designed, the participants were selected, and the questions were posed. Only the P1 code (Proportional, risk ratio reasoning with interior cells) or the P3 code (Proportional, odds ratio reasoning) were pertinent to most of the tasks.

As I analyzed each of the interviews, I worked with the multi-view video, the transcript, the scanned student work, and my notes. It is not possible to re-create every aspect of the interview, but working with these multiple data sources across interview participants for different tasks within the interview is a type of triangulation (Denzin &

Lincoln, 2005; Patton, 2005) that helped me to develop a comprehensive understanding of the participant's thinking. I considered the reasoning of the participants to make refinements to the P1 and P3 framework codes and also to identify other phenomena. When I observed a more fine-grained way a student was reasoning, I created a new code and a description. As the analysis ensued, I created more codes, made analytic notes and modified, combined, or deleted some codes. I summarized each interview with thick, rich, narrative descriptions and once I had a set of codes for a modified framework, I created a summary identifying each code, its description, and how it applied to each task for each participant. I reviewed this information with another researcher, made additional changes, and then reviewed them in relation to each task and each participant. For this review, I considered the participants in the opposite direction, which allowed me to become more familiar with their work and resulted in a refinement of the codes. The resulting modified framework is presented in Table 12.

Table 12

Modified Framework for Reasoning with Contingency Tables

Code	Name	Description and features
N1CMF	No interior cells are used, compared marginal frequencies	None of the interior cells in the table are used, rather the marginal frequencies are used to compare with one another
L2CP	Localist 2 interior cells comparing proportions	Two of the joint frequencies or conditional relative frequencies for one condition in the table are compared with one another where one is less than, equal to, or greater than the other
P3E	Odds approach noticing equality	The joint frequencies in the table are used to compare with one another where the equality of values is recognized

Code	Name	Description and features
P3AN(MP)	Odds approach noticing all and none	The joint frequencies in the table are used to compare with one another either within rows or columns where all and none (or close to all and none) values are recognized and (MP) is added when the mosaic plot is used in addition to the numbers in the contingency table
P3R(MP)	Odds approach considering ratios	Considers ratios of joint frequencies to compare with one another where one is less than, equal to, or greater than the other and (MP) is added when the mosaic plot is used in addition to the numbers in the contingency table
P1SW	Risk approach noticing equal marginal frequencies and comparing whole numbers	Compares two of the joint frequencies, noting their marginal frequencies are equal for complete tables and selecting equal numbers for marginal frequencies for incomplete tables
P1SF	Risk approach noticing equal marginal frequencies and comparing fractions	Compares two of the conditional relative frequencies, noting the marginal frequencies are the same for complete tables. Thus, the numerators of the fractions can be compared
P1ED	Risk approach, using equations resulting in decimals	Uses division equations to consider conditional relative frequencies resulting in a decimal to compare, noting one is less than, equal to, or greater than the other

Code	Name	Description and features
P1FD	Risk approach, using fractions resulting in decimals	Considers conditional relative frequencies as fractions and uses decimal notation to compare, where one is less than, equal to, or greater than the other
P1BF(MP)	Risk approach, using a benchmark fraction to compare	Considers conditional relative frequencies as fractions and uses a benchmark fraction of ½ where one is less than ½ and the other is greater than ½ and (MP) is added when the mosaic plot is used in addition to the numbers in the contingency table
P1F(MP)	Risk approach, using fractions	Considers conditional relative frequencies as fractions where one is less than, equal to, or greater than the other and (MP) is added when the mosaic plot is used in addition to the numbers in the contingency table
P1SU	Risk approach, using marginal frequencies to scale up the smaller condition	Considers the different marginal frequencies and scales up the row with the lesser value to equal the greater value then compares the joint frequencies noting one is less than, equal to, or greater than the other
P1MP	Risk approach, using only the mosaic plot	Considers the mosaic plot and compares either the length or area of the blocks where one is less than, equal to, or greater than the other
P2	Proportional, risk ratio reasoning with interior and exterior cells	All four cells in the table are used to decide about independence or association and multiplicative reasoning is used. Proportional reasoning compares risk (part to whole ratios) and **compares one conditional relative frequency to a marginal relative frequency**, using both interior and exterior cells.
WD	Wrong Direction	Conditions using the wrong direction (e.g., rows instead of columns)

Researcher Bias

Researcher subjectivities can also be considered a limitation in qualitative research. I came to this study with my own experiences and preferred strategies I have developed when solving similar problems. Personally, I have struggled with keeping up with multiple components of things, whether they are physical items or quantities in a math or statistics problem. When it comes to proportional reasoning, I have struggled at times to reason both accurately and quickly, often having to break apart constituent components for accuracy or making mistakes when trying to reason quickly. I often feel like this is something I should have learned in middle school because it is viewed as basic and unsophisticated mathematics. However, I have seen students of many levels, including those studying advanced mathematics, struggle with proportional reasoning as well as categorical data association with contingency tables. This limitation may have potentially hindered me from asking good, unplanned, probing questions during the interviews. Because of this I made sure to include alternative solution methods and possible probing questions in the interview protocols. I reviewed each protocol in detail before each interview so that I was grounded in the context and numbers for the included tasks as the participants explored the association of categorical variables with contingency tables.

As I summarized the data, I tried to avoid bias. I used separate documents for descriptions of what occurred versus my opinions and thoughts. When creating lesson graphs (see Figure 11), I had a separate column for "Student Action," where I summarized what the participants did and "Notes," where I included my observations and ideas.

Figure 11

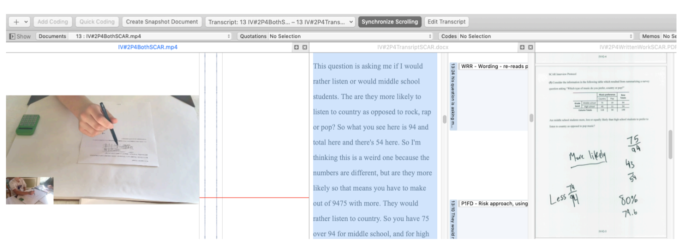
Lesson Graph Example

IV#2P3SCAR Lesson Graph			Zander
Time	***Student Action***	***Notes***	***Written Work***
Task 1 (MS/HS Survey happy & rich, all=)			
2:16:00	Re-reads question	focuses on interior cells	equally likely
Equally	Compares middle school numbers to one another		
	Then compared high school students to one another		
Task 2 (MS/HS Survey happy & healthy, 0's more)			

As I analyzed the research data, I recognized I had certain biases, some I was not even aware of, and I made efforts to manage these. Coding is an interpretive act, not an exact science (Saldaña, 2015), and the way I interpreted the data from the interviews was influenced by my past experiences and how I choose to look at the data. When applying codes from the framework, I used the videotaped interviews, transcripts, and participants' work (see Figure 12) and did not consider my field notes or observations made in lesson graphs. When I questioned which code to apply, I re-watched the video, re-read the transcript, and reviewed the student work to base decisions on the data as much as possible.

Figure 12

Coding Example (from ATLAS.ti)



CHAPTER 4

FINDINGS

This study focuses on 20 think-aloud interviews for five different participants. In this chapter, I first discuss each of the five participants and their reasoning across the four interviews they each completed. Next, I present my findings and use my research questions to frame the results:

1. In what ways do students reason about (in)dependence of categorical variables when using contingency tables?
2. In what ways do students use mosaic plots to reason about (in)dependence of categorical variables when using contingency tables?

I use a cross-case analysis, which includes application of my framework, where I use the second (IV#2) and fourth (IV#4) interviews with complete and incomplete contingency tables to address the first research question and the third (IV#3) and fifth (IV#5) interviews with contingency tables and corresponding mosaic plots to address the second research question. Throughout the analysis, I make connections to the literature, noting where these findings support, differ from, and extend the existing research. Additionally, I address challenges the participants had with wording and direction when working with contingency tables. Finally, I conclude with a summative discussion of the findings.

Summary of Participants

The five participants whose work is considered for this study were middle and high school boys from a suburban area who attended public schools. The interviews were conducted outside of regular school hours either after school, on weekends, or over holiday breaks. Here, I provide a brief summary of each participant considered for this study (Jamie, S1; Zander S2; Sydney, S3; Hayden, S4; Jessie, S5).

Jamie, the first student in this study (S1), was an eighth-grade student in accelerated Algebra I/Geometry, which covers the entirety of on-level ninth-grade math and the first half of 10th-grade math. He was taking science for high school credit (physical science) and identified social studies and Spanish as his favorite subjects. Jamie mentioned he had seen contingency tables for “only a few weeks” in seventh grade but had not seen a mosaic plot before. He generally used fraction notation and converted to decimal notation for comparison when the denominators were not equal and a mosaic plot was not provided. Jamie recognized he could determine (in)dependence with the mosaic plot alone or in conjunction with any 2x2 contingency table. When reasoning with a mosaic plot, he maintained a focus on comparing the “percentage rate” for the conditioned variable.

Zander, the second student in this study (S2), was also an eighth-grade student in accelerated Algebra I/Geometry math class and taking high school science. He identified social studies as his favorite subject and was eager and excited to participate in the interviews. He used a variety of markers to color code when working through the tasks and used whole-to-part, part-to-whole and part-to-part ratios, and equations throughout the tasks. Zander found a mosaic plot to be useful and solved problems using it alone and in conjunction with a contingency table. The mosaic plot helped Zander solve a problem with an incomplete contingency table that he was unable to solve without the mosaic plot (IV#4, Task 4(a) and IV#5, Task 4(a)). His reasoning with mosaic plots was flexible and appropriate. Sometimes he used ratio reasoning (P1) and other times he used odds reasoning (P3); sometimes he used a measurement of area, and other times he used a measurement of distance.

Sydney, Zander's brother and the third student in this study (S3), was a ninth-grade student in Honors Geometry and on-level science (physics). He identified chorus and weight training as his favorite subjects. Sydney struggled the most with the wording of the questions, requesting to skip one of the tasks in the first and third interviews. He thought he had seen a mosaic plot before from a substitute teacher but could not remember specifics. Sydney admitted to not liking to draw or use pictures in math and preferred using the numbers in the contingency tables. Sydney was not clear that a mosaic plot could be used to determine (in)dependence for all 2x2 contingency tables. He was the only student able to solve the task with an incomplete contingency table and no mosaic plot (IV#4, Task 4(a)) where both marginal frequencies were provided and a situation of independence was given. However, he needed assistance when the same problem was presented with a mosaic plot (IV#5, Task 4(a)).

Hayden, the fourth student in this study (S4), was a ninth-grade student in Accelerated Geometry B/Algebra II which covers on-level math typically taught in the second half of 10th grade and the entire 11th grade. Hayden played competitive basketball and identified Spanish and Advanced Placement Human Geography as his favorite classes. He systematically worked through tasks expeditiously, using equivalent fractions and admitting that he did not like to estimate but rather preferred to be exact in mathematics so he could get the right answer. This was evident in his requiring exact equivalence to conclude "equally likely" as well as his preference for the numbers in the contingency tables versus estimates he could derive from the mosaic plots. Hayden found the mosaic plots to be more useful when the marginal frequencies were equal and primarily focused on the area of the blocks rather than the linear distance.

Jessie, the fifth student in this study (S5), was a seventh-grade student in Math 7/8, which covers on-level math typically taught in the last half of seventh grade and the entire eighth-grade year. He identified his favorite subjects as math, social studies, and English/literature and mentioned being afraid of getting the wrong answer due to his experiences with taking tests in math class. Similar to Jamie and Zander, Jessie found the mosaic plots to be useful and readily began using them alone to reason about (in)dependence.

All five of these participants were tracked into a higher math class. For example, Algebra I is taught in ninth grade for on-level math, and each of these participants either completed or was on track to complete Algebra I by the end of eighth grade. This was not usual for their schools, although the Algebra I participation rate for their middle school was 27%, as compared with the state rate of 19% (greatschools.org). They were all tracked into a higher-level math beginning in middle school, and four of the five participants were in a math class with an accelerated pace, covering a year and a half of material in comparison with typical classes. Because the interviews took place outside of school, participants were not missing any instruction. Thus, the students' motives for participation in my study did not include missing class. This factor may be attributed to all five participants being in advanced math courses. An interesting note is that these participants mentioned they wished their teachers went into more depth when explaining things in their mathematics classes.

Participants' Reasoning with Contingency Tables

My first research question aims to understand how students reason about (in)dependence with contingency tables. This includes contingency tables that are both complete and incomplete. Providing students with an incomplete contingency table and asking them to fill in data that will create a situation with (in)dependence requires what Inhelder and Piaget (1958) referred to as reversibility where the result is given and students coordinate actions to come back to a starting state. Interview #2 included complete contingency tables, which have been the focus of past studies. Interview #4 included incomplete contingency tables, which require reversibility in thought and have been absent from the literature.

Reasoning with Complete Contingency Tables (IV#2)

The tasks for IV#2 along with answers and goals of the interview are included in Appendix C. The goal of the second interview was to begin to understand how students who can apply their proportional and probabilistic reasoning in context work with categorical variables in contingency tables to determine (in) dependence. IV#2 included eleven tasks, with three tasks having two parts, (a) and (b).

Tasks with Complete Contingency Tables (IV#2)

Tasks in IV#2 begin with a context of a survey for middle school and high school students where the first 3 questions are about lifestyle and the next 4 questions are about music preference. The last 4 questions include contexts of flu shot type and flu status, drug types and disease status, handedness and height, and cereal and shelf location. The questions asked consistently focused on the category in the first column and required students to condition on the rows. The exception to this was in part (b) of problems 9, 10, and 11 where questions required conditioning on the columns. The first 9 problems asked questions that requested a selection between more likely, less likely, or equally likely, whereas the last two questions asked for a more general comparison in context. Table 13 summarizes the tasks along with some different aspects of the problems that were discussed in Chapter 2 (see Table 5, p. 37).

Table 13

Task Summary for Complete Contingency Tables (IV#2)

Task	Context Description	Correct Answer	Condition Direction	Row Marginal Frequencies	Question phrasing M/ L/E	Other Features
1	MS/HS Happy/Rich	Equally Likely	Row	Same	Yes	Joint frequencies all equal
2	MS/HS Happy/Healthy	More Likely	Row	Same	Yes	Joint frequencies have 0's on diagonal
3	MS/HS Healthy/Rich	Less Likely	Row	Same	Yes	Joint frequencies have 0's on diagonal
4	MS/HS Rap/Rock	Equally Likely	Row	Same	Yes	
5	MS/HS Country/Pop	Equally Likely	Row	Different	Yes	as close to equal proportions as possible
6	MS/HS Country/Rock	More Likely	Row	Same	Yes	
7	MS/HS Rap/Country	More Likely	Row	Different	Yes	Proportions on either side of 1/2
8	Flu Shot/Nasal Mist Flu/No Flu	More Likely	Row	Same	Yes	
9	Drug A/Drug B Disease/No Disease	More Likely	Row	Different	Yes	Close proportions, different sample sizes
9b	Drug A/Drug B Disease/No Disease	More Likely	Column	Different	Yes	
10	Righty/Lefty Taller/Shorter	Equally Likely	Row	Different	No	
10b	Righty/Lefty Taller/Shorter	Equally Likely	Column	Different	No	
11	Kids'/Adults' Cereal Upper/Lower Shelf	Less Likely	Row	Different	No	
11b	Kids'/Adults' Cereal Upper/Lower Shelf	Less Likely	Column	Different	No	

Applying the Framework to IV#2 Data

Participants generally answered the first 3 tasks using the 4 interior cells (joint frequencies) and comparing them to one another. This is similar to using an odds ratio type of reasoning (P3 from the initial framework) as opposed to a relative risk type of reasoning (P1 or P2 from the initial framework). For subsequent tasks, participants primarily considered a risk type of reasoning using both joint frequencies and marginal frequencies to compare one conditional relative frequency to another (P1). Table 14 summarizes the codes assigned for each participant across all of the tasks in IV#2.

Table 14

Framework Application for Complete Contingency Tables (IV#2)

Task	Jamie(P1)	Zander (P3)	Sydney (P4)	Hayden (P6)	Jessie (P8)
1	P3E	P3E	P1SW, P3E	P3E, P1SF	P3E, P1SW
2	P3AN	P3AN	P3AN	P1SF	P3AN
3	P3AN	P3AN	P3AN	P1SF	P3AN
4	P1SF	P1SW	P3R	P1SF	P3R
5	P1FD	P1ED	P1FD	P1F, P1FD	P1SU, P1F
6	P1SF	P1SW	P1FD	P1SF	P1SW
7	P1BF	P1ED	P1FD	P1BF	P1SU
8	P1SF	P1SW	P1FD, P1SW	P1SF	P1SW
9	P1FD	P1ED	P1FD	P1F	P1SU
9b	omitted	*L2CP, P1ED, WD*	start P1FD*	P1F	omitted
10	P1FD	*P1ED, WD,* P1ED	P1FD	P1F	P1SU
10b	*N1CMF, L2CP,* P1FD	*L2CP,* P1ED	P1FD	P1F	omitted
11	P1FD	omitted	P1FD	P1F	omitted
11b	P1FD, WD**	omitted	*L2CP,* P1FD	P1F	omitted

* Student started the task but requested to skip it
** Task was only read aloud and was not given in a written form

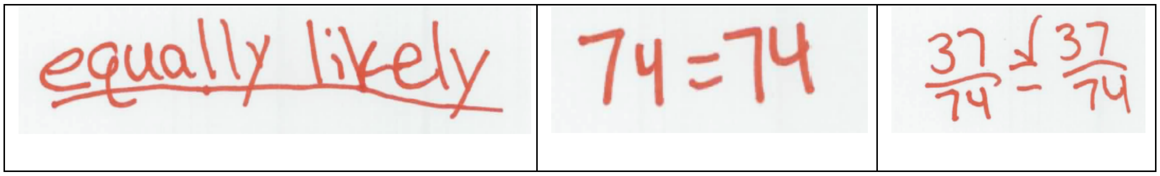
The first seven tasks included a context where middle and high school students were surveyed. For the first task, almost all participants initially noticed that the four interior cells were all equal, which is reflected by the P3E code. Sydney was the only participant who, to begin with, noted the marginal frequencies for middle and high school students were the same and then compared the joint frequencies; thus I coded his work for this task as P1SW. He said, "There's the same amount of students for each, and they're all the same number." This is a different reasoning than the incorrect strategy (S10) that Batanero described as follows:

> S10 – Use of marginal frequencies. Some students consider the problem impossible to solve because of the difference in marginal frequencies in different rows or columns in the table.

In this case, Sydney noticed the marginal frequencies were equal, and rather than use this as a criterion to solve the problem, he used the equal marginal frequencies to simplify the calculations for the problem. He recognized that the equal marginal frequencies meant he could compare the whole number joint frequencies rather than having to compare proportions. Similarly, Hayden also recognized the equal marginal frequencies and furthermore used those as denominators to compare row conditional relative frequencies (RCRFs) that he represented as fractions (see Figure 13). Hayden kept this P1SF strategy throughout the remainder of the tasks in IV#2, adjusting it when marginal frequencies were not the same (P1F) and using benchmark fractions for comparison (P1BF).

Figure 13

Hayden's P1SF Work for IV#2, Task 1

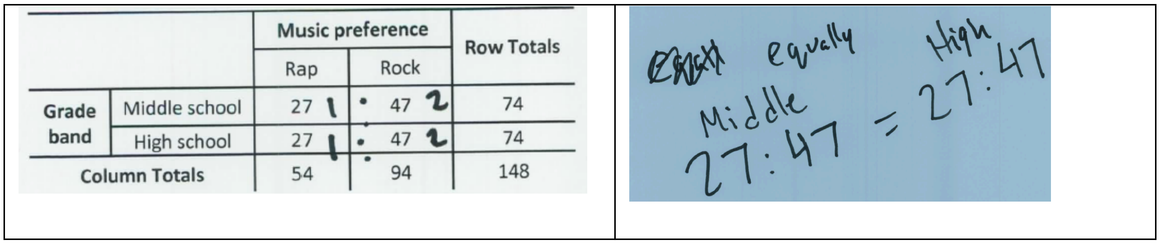


On the second and third tasks that presented problems with complete dependence, most participants recognized the zeroes on the diagonals of the interior cells as they noted the all or none values for the survey response of life preferences (P3AN). For example, in task 2, zero middle school students preferred to be healthy and zero high school students preferred to be happy. Beginning with the fourth task, three of the five participants started to use more of a part-to-whole approach (P1), similar to what Hayden was using where they focused on risk rather than odds. They compared RCRFs using fractions (P1F) or recognized the row marginal frequencies (RMFs) were equal and simply compared the whole number joint frequencies (P1SW). On the other hand, Sydney and Jessie, rather than using part-to-whole ratios (risk), compared the part-to-part ratios (odds) for middle and high school students for task 4. For example, Sydney's work is shown in Figure 14 where he considered the ratio of the categories of the music preference variable, rap to rock as he said:

> It'd be like, one, it's like, I know they're not, it's not one to two, but it'd be like one to two is one to two. Like the ratio. I know it's not; this isn't 54 but like say, this is just one, this is two. It'd be like that. And then this is also one. And this is also two. So, they're equal.

Figure 14

Sydney's Work for IV#2, Task 4



		Music preference		Row Totals
		Rap	Rock	
Grade band	Middle school	27	47	74
	High school	27	47	74
Column Totals		54	94	148

Risk and odds are easily conflated (Ranganathan et al., 2015), and understanding their difference is important for understanding more advanced statistics (e.g., odds ratio, risk ratio, logistic regression, etc.). Whereas the participants' work does not indicate whether they clearly understood the difference between odds and risk, it does demonstrate they could use both types of reasoning.

Beginning with task 5, all students began comparing conditional relative frequencies (P1). This type of reasoning likely occurred because the questions were worded in a way that asked them to consider particular categories of each variable in comparison with one another. The first four tasks were also worded this way, but the numerical features of these tasks may have encouraged students to use an odds type of approach. The contingency table for task 5 (see Table 15) had a unique design feature beyond the unequal RMFs where the RCRFs were approximately equal ($75/94 \cong 43/54$, or $0.7978 \cong 0.7962$). Given the RMFs are unequal but do not have a common denominator other than 2, there are limited choices for a scenario where the RCRFs are exactly equal. Fixing the numbers for the high school students (43 prefer country and 11 prefer pop), the number of middle school students represents a situation where the CCRFs are as close to equal as possible because, as two of the students recognized, there cannot be a partial student. Moving one middle school student from preferring country to pop results in another approximately, but less equally likely scenario (74/94 ≅ 43/54, or 0.7872 ≅ 0.7962).

Table 15

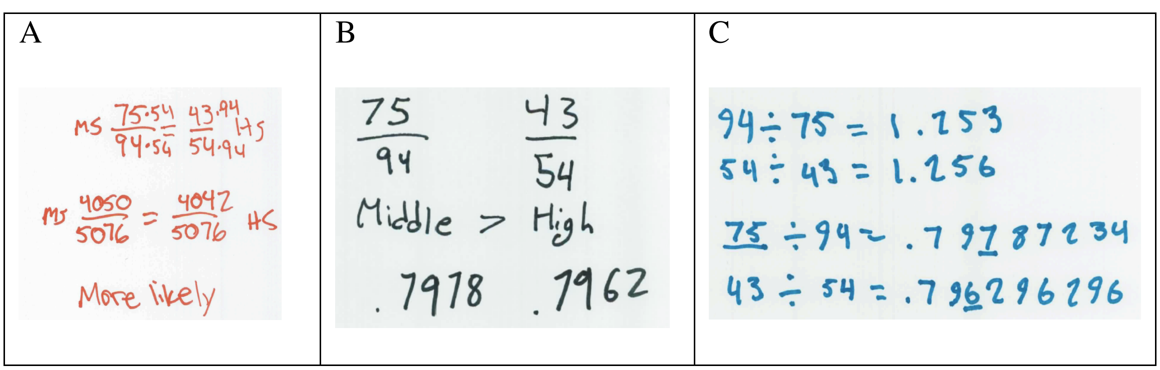
Contingency Table for Country and Pop Music Preferences (IV#2, Task 5)

		Music preference		Row Totals
		Country	Pop	
Grade band	Middle school	75	19	94
	High school	43	11	54
Column Totals		118	30	148

All students used proportional reasoning for this task where Hayden maintained his approach using strictly fractions, Jamie and Sydney converted fractions to decimals, and Zander used equations to compute decimal proportions (see Figure 15). Jamie noted he did not continue to use fractions because of their denominators, “So then we would simplify to see which one is bigger. But it looks like they’re not divisible by each other.”

Figure 15

Participant Solutions Using Proportional Reasoning (IV#2, Task 5)



Note. Panel A: Hayden’s work (P1F). Panel B: Jamie’s work (P1FD). Panel C: Zander’s work (P1ED).

Jessie used a different approach where he noticed that the row marginal frequencies were different, "…upon seeing that there are more middle schoolers than high schoolers. You have to take that into consideration. Because you can't compare even groups anymore." He suggested adding a little bit more to the smaller group and to determine how much more, and he did this by adding the same proportion. He computed a multiplier by dividing the larger row marginal frequency for middle school by the smaller row marginal frequency for high school (94 ÷ 54 = 1.74). He then verified his work (1.74 x 54 = 93.96 ↑ = 94) and then applied the 1.74 multiplier he found to the joint frequency for the middle school students who prefer country (43 x 1.74 = 74.82 ↑ = 75). Because the resulting product rounded to the same number for high school students who prefer country, he concluded they were equally likely. Without prompting, he proposed a different strategy of "a fraction way," comparing the row conditional relative frequencies, but he mentioned you cannot compare fractions unless the denominators are the same. Since the denominators did not have common factors, he did not continue with this approach, but rather connected this to his earlier strategy. Jessie mentioned you would still have to compute the multiplier he found earlier, mentioning that you have to multiply the numerator and the denominator by the same value to keep the ratio the same and to compare them to one another. He also noted the resulting greater joint frequencies will add to the greater marginal frequency. Jessie explained that if he applied the multiplier and got different joint frequency numbers in comparison with the other row, they would no longer be equally likely, but either more or less likely.

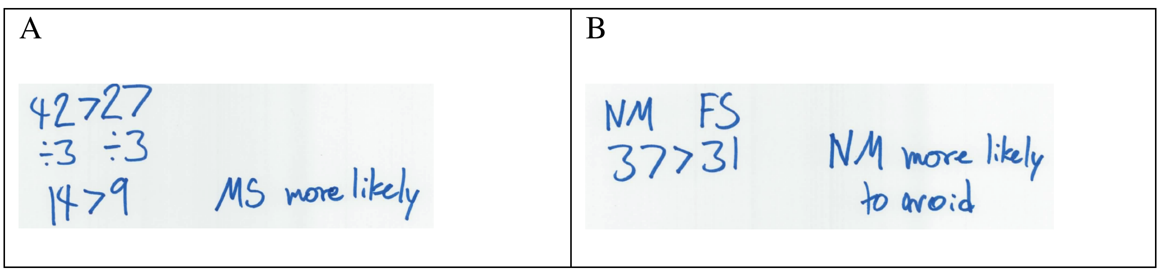
Whereas none of the students claimed "equally likely" to be the clear and definitive answer, they differed in their willingness to consider it as a possible solution.

Jamie noted they were both equal to ".79 something" (he recalculated to get more specificity) and were "very, very, very" slightly different. When asked to choose between equally likely and more likely, he mentioned if he were looking at "statistics directly" he would choose more likely, but in general they were equally likely. Zander also noted the proportions were "pretty much close," referenced the place value where they differed and claimed middle school students have a "little greater chance." Similarly, Sydney stated, "Honestly, though it would probably be about equally likely." But like Hayden, he concluded the given problem was "more likely" and if the 75 were changed to 74 it would be "less likely." Hayden was on the other end of the spectrum and claimed that because the numerators in his fractions were different, it was not "equally likely". When asked if he could use a different approach, he computed proportions and maintained that middle schoolers were more likely than high schoolers to prefer country music because .798 is greater than .796 and if one middle school student were moved from the country category, they would be less likely. My interpretation of the participants' actions in this problem highlights the difference between statistical and mathematical thinking (Cobb & Moore, 1997). Statistically we want students to recognize there is no evidence to conclude the two situations with 74 or 75 middle school students are not equally likely; however, mathematically we want students to be able to compare two fractions as numbers and recognize even a small difference. Even a small difference of one or two percentage points or moving one or two more students would not provide strong evidence to support they are not equally likely.

For the remaining tasks, participants continued to use strategies similar to what they used for task 5; however, they often used features of the problem to allow for simpler calculations. This included comparing whole numbers when the marginal frequencies were the same (P1SW) and using a benchmark fraction of 1/2 (P1BF) if one of the RCRFs was greater than 1/2 and the other was less than 1/2. For example, Jessie (see Figure 16) noticed there was the same number of middle and high school students in task 6 and the same number of people who received the flu shot and nasal mist in task 8, so he simply compared the whole number joint frequencies. Jessie's work here and other students' work using this strategy coincides with past studies of student reasoning where an alternative strategy is considered to avoid fractions (Karplus et al., 1983b).

Figure 16

Jessie's Work With Equal Marginal Frequencies Coded as P1SW



Note. Panel A: IV#2, task 6 work. Panel B: IV#2, task 8 work.

In task 7, Jamie and Hayden used a benchmark fraction of 1/2 to compare the fractions (P1BF). They both doubled the numerator and compared it to the denominator, noting that if the doubled numerator was greater than the denominator, then the fraction was greater than 1/2 and if the doubled numerator was less than the denominator, the fraction was less than one half. Researchers found this transitive approach was used by students who successfully reasoned when comparing fractions (Clarke & Roche, 2009). This supports the determination that I made about these students' proportional reasoning abilities and also indicates this type of question might be used as a type of formative assessment for proportional reasoning.

In these instances, it does not seem that students' strategies are deficient when they recognize the marginal frequencies are equal and compare two joint frequencies or when they do not explicitly compare proportions. On the contrary, the fact that they simplify calculations demonstrates numerical fluency and flexible thinking.

Reasoning with Incomplete Contingency Tables (Interview 4)

The tasks for IV#4 along with answers and goals of the interview are included in Appendix E. The goal of the fourth interview was to understand how students reason about (in)dependence with contingency tables that are incomplete. IV#4 included five tasks, each with three parts asking for each of the situations less likely, equally likely, and less likely.

Tasks with Incomplete Contingency Tables (IV#4)

Tasks in IV#4 had similar contexts as those in IV#2 and IV#3 and once again, different aspects of contingency tables were considered. Each of the tasks included 3 parts, (a), (b), and (c), that each used the same contingency table but asked the students to complete the missing entries in the table to represent different situations of less likely, equally likely, and more likely. Like the earlier tasks, the questions focus on comparing categories where one row and column are used as the basis for comparison. The questions started with fewer values given, allowing for a wider variety of answers, and progressed to more restrictive and challenging combinations of values. The first task had a context of a survey for children and adults and their pet preference, a dog or cat. The remaining tasks included contexts of exercise frequency and sleeping problems, drug types and disease status, and cereal type and shelf location. The first four questions consistently asked students to complete the values in the table to represent situations where the first or second category of the row variable was more, less, or equally likely than the other category of the row variable. The fifth task asked for situations where one of the categories of the column variable was more, less, or equally likely than the other category of the column variable, in essence asking for CCRFs to be compared as opposed to RCRFs. The order of more, less, and equally likely was different across the tasks. Table 16 summarizes the tasks along with some different aspects of the problems.

Table 16

Task Summary for Incomplete Contingency Tables (IV#4)

Task	Context Description	Completed Cells	Direction and Base Row/Column	Row Marginal Frequencies	Order (a)/(b)/(c)
1	Adults/Children Dog/Cat	Total	RCRF, 1, 2	Undefined	More/Less/Equal
2	Exercise Frequently/Seldom Sleeping Problems/No Problems	Row Marginal Frequencies	RCRF, 1, 2	Different	Less/Equal/More
3	Drug A/B Cured/Disease	Row Marginal Frequencies and one Joint Frequency	RCRF, 2,1	Different	Equal/Less/More
4	Cereal Name/Store Brand Upper/Lower Shelf	All Marginal Frequencies	RCRF, 1, 2	Different	Equal/More/Less
5	Drug C/D Cured/Disease	Row Marginal Frequencies and one Joint Frequency	CCRF, 1, 1	Different	Equal/Less/More

Applying the Framework to IV#4 Data

Participants generally answered tasks using the same reasoning and strategies they had used in IV#2. One student, Jessie used a mosaic plot, which he was introduced to in IV#3. All participants used whole number comparison when the marginal frequencies were equal and benchmark fractions to select numbers. Table 17 summarizes the codes assigned for each participant across all of the tasks in IV#2.

Table 17

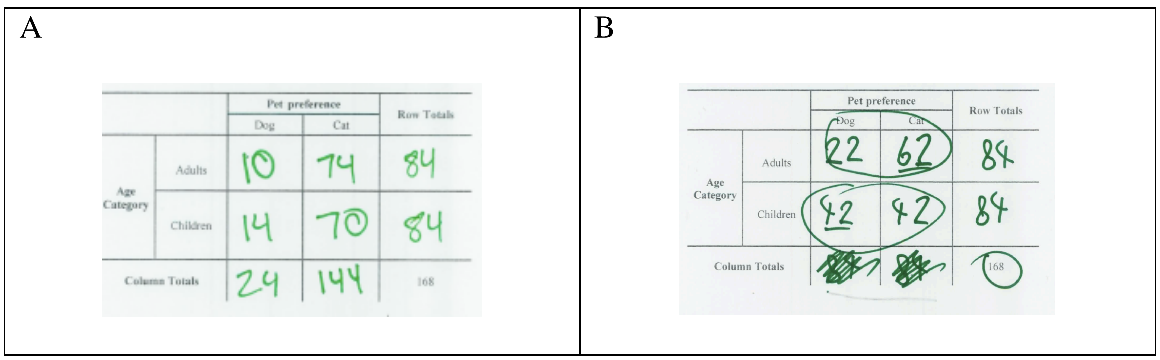
Framework Application for Incomplete Contingency Tables (IV#4)

Task	Jamie(P1)	Zander (P3)	Sydney (P4)	Hayden (P6)	Jessie (P8)
1(a)	P1SW, P3AN, P1BF	P3AN, P1FD	P1SW	P1SW, P3AN, P1F	P1SW, P1BF
1(b)	P1SW, P1BF	P1SW, P1FD	P1BF	P1SW	P1SNR, P3AN
1(c)	P1SW, P1E	P3E, P1SW	P1BF	P1SW	P3E, P1BF, P1SW
2(a)	P3AN, P1BF, P1FD	P1BF	Skip	P1BF, P1F	P3AN
2(b)	P1BF	P1BF, P1ED	P1BF	P1BF, P1F	P1BF
2(c)	P1BF	P1BF	Skip	P1BF, P1F	P1BF, P1SU
3(a)	P1FD	P1ED	P1FD	P1F	P1SU
3(b)	P3AN	P1FD	P1FD	P1BF, P1F	P1AN
3(c)	P3AN	P1FD	P1FD, P1AN	P1BF	P1AN
4(a)	Skip	*Start P1FD*	P2	*Start P1F*	StartP1F
4(b)	P3AN	Omitted	*L2CP*	P1BF, P1F	P3AN
4(c)	P3AN	*P1FD*	*L2CP*	P1BF, P1F	P3AN
5(a)	omitted	omitted	P1FD	omitted	P1SU
5(b)	omitted	omitted	omitted	omitted	omitted
5(c)	omitted	omitted	omitted	omitted	omitted

The first task was the only one that allowed students to select both the row and column marginal frequencies. Most of the participants began with making the row marginal frequencies equal (P1SW), as Jamie said, “to make it easier.” The exception was Zander, who preferred to use a variety of approaches to push his thinking. He placed one observation for adults, which forced this category to an all and none distribution (P3AN). He later noted this did not make it a very reliable survey. For parts (b) and (c), he switched to equal RMFs. When students considered how to divide the RMFs between the column categories for part (a), they either chose to put most of the adult observations in the prefer cats category and then chose a number that was smaller for the children who prefer cats (see Figure 17A) or they split one of the RMFs equally and adjusted the RMF for the other row (see Figure 17B).

Figure 17

Solutions Partitioning Equal Marginal Frequencies (IV#4, Task 2(a))



Note. Panel A: Hayden’s P1SW work for 1(a). Panel B: Jessie’s P1BF work for 1(a).

Additionally, for part (a) of task 1, students used reasoning similar to what they used in IV#2. For example, Hayden continued to use equivalent fractions and Jessie continued to scale the values in one row to compare to the other. Jessie also mentioned his preconceived notion aligned with adults being more likely than children to prefer cats as opposed to dogs. This preconception about context surfaced for other students across the interviews, but it did not seem to interfere with their reasoning. As Jamie said in a later problem, he was used to putting his real-world understandings aside to consider the mathematics in a given problem.

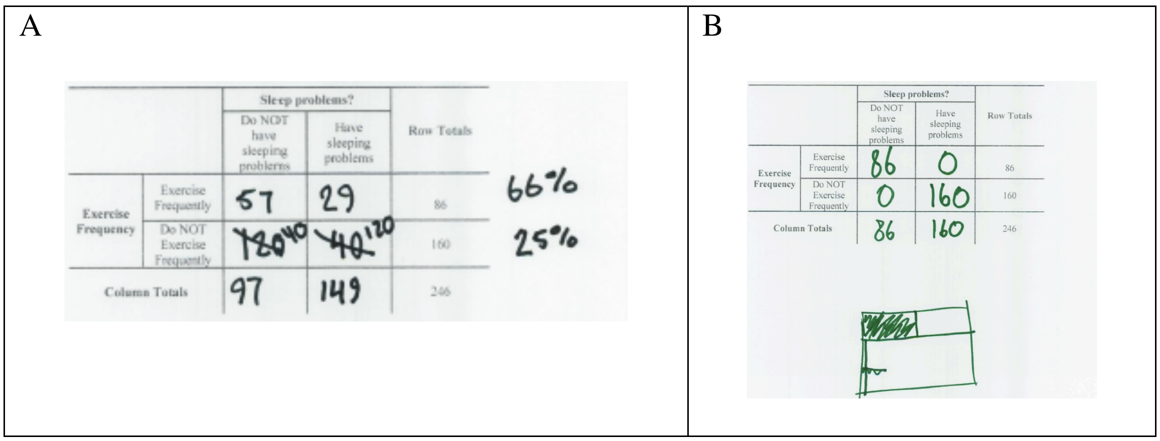
Task 2 asked students to create a situation where those who exercise frequently are less, equally, or more likely than those who do not exercise frequently to have sleeping problems, as opposed to not have sleeping problems. This task focuses on the second column (have sleeping problems) and required the students to condition on and compare the categories of the row variable of exercise frequency. All students used a benchmark fraction of 1/2 at some point in this task. One student, Sydney requested to skip this task, but later completed part (b). Later, I discuss some possible reasons for this.

Part (a) of task 2 asked for a less likely situation and most of the students either split one of the row marginal frequencies equally and adjusted the other one (see Figure 18A) or chose an all and none approach (see Figure 18B). Zander used benchmark fractions of 2/3 and 1/4 rather than 1/2 because he wanted to use different proportions than he previously used. It is interesting to note Jessie did not consistently use the same approach he used in task 2, which demonstrated his flexibility in reasoning. On the other hand, some of the other students' strategies, like equivalent fractions and scaling up remained consistent throughout the tasks. Participants recognized a range of numbers that could work for this situation and when determining and explaining this range, Jessie drew a mosaic plot.

Part (b) requested an equally likely situation and all participants halved the given marginal frequencies (P1BF) and although some initially thought this was the only solution, they all recognized a range of solutions were possible. Similarly, students swapped the values in rows for part (c), noticing the inverse relationship between the more and less likely situations. Students gave reasons why their preconceived ideas about the context of this task might cause them to think one way or another (e.g., more fit bodies rest better or more exercise can cause more injuries). When asked to create an equally likely scenario, Jamie did not think his pre-conceived ideas interfered because math problems in school are often not aligned with the real-world.

Figure 18

Solutions Partitioning Unequal Marginal Frequencies (IV#4, Task 2(a))



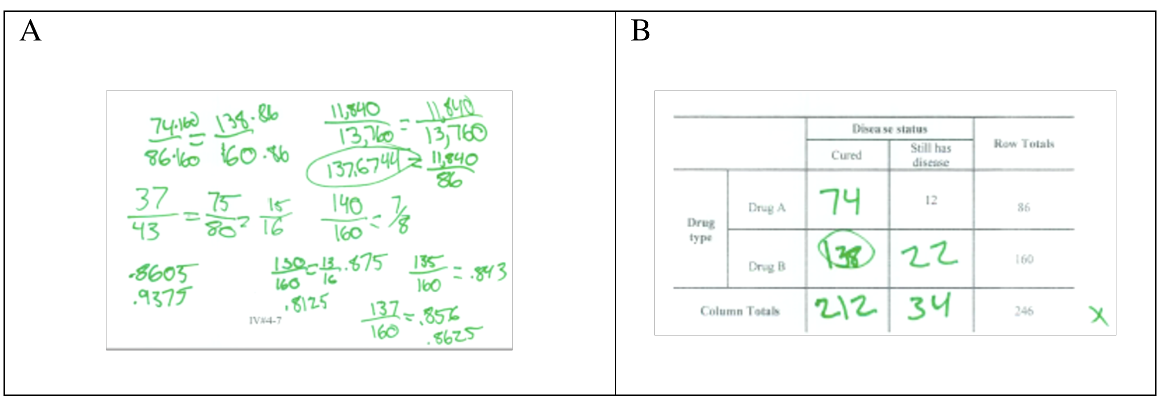
Note: Panel A: Zander's P1BF work for 2(a) with 2/3 and 1/4. Panel B: Jessie's P3AN work for 2(a).

Task 3 was more restrictive in that the value of one of the interior cells (12 patients who received Drug A were not cured) was provided in addition to row marginal frequencies. For part (a), all participants noted the additive relationship of the joint frequencies and calculated the 74 patients who were cured and received Drug A. Next, most participants calculated the RCRF (74/86 = 86%). Jessie, on the other hand, used his scaling-up strategy and found the multiplier 1.86 to scale the smaller row by dividing the marginal frequencies (160 ÷ 86). He was going to multiply one of the joint frequencies of the smaller row (12 x1.86), but he chose not to do that as he was unsure whether it would provide the correct answer. Jamie felt it was difficult to transition to a think-aloud interview because of the frequent tests he took in school. In addition to the reminder about not being interested in right or wrong answers that I read at the beginning of each interview, we had discussions throughout the interviews about how this was not a test and all of his thoughts were important. As he resumed his work, he multiplied the other joint frequency for adults (74 x 1.86) and then completed the chart. In the end he was fairly confident in his answer.

All participants were able to determine the answer to task 3(a), which demonstrated reversibility. However, Hayden struggled since he could not find an exact equivalent fraction (see Figure 19). He identified that 74/86 was 0.8605, determined that 137 and 138 were the closest possible answers, and initially claimed there was not an answer. After he considered the closest answer of 138 and worked with the later parts of the problem, grappling with whether 138 was actually more likely, Hayden decided 138 was close enough to be equally likely.

Figure 19

Hayden's Work with Equivalent Fractions (IV#4, Task 3(a))



Note: Panel A: Hayden's initial work for 3(a). Panel B: Hayden's final answer for 3(a).

Part (b) asked for less likely and students used benchmark fractions, an all and none approach, or selected a lower percentage and multiplied it by the marginal frequency of

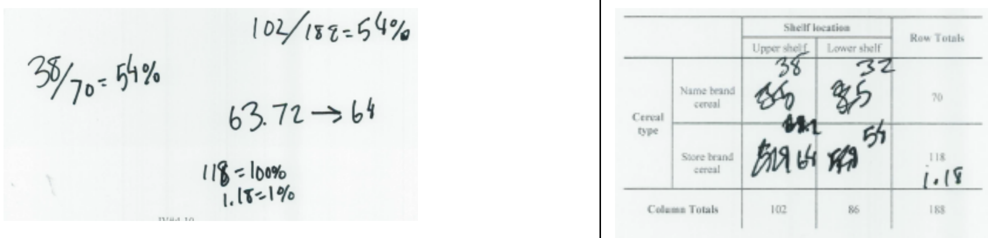
160. Students were able to recognize a range of possible answers and claimed the less likely situation would hold up to the point where the numbers were the same as the equally likely situation in part (a). For part (c), students either used the same approach they used in part (b), or they used the information from the equally likely situation in part (a) as a guide.

Task 4 included all marginal frequencies and part (a), which asked for an equally likely situation, proved to be the most challenging. Sydney was the only participant to create a solution and unlike the other participants, he had just worked with a problem asking for an equally likely situation (2(b)). This may have been a contributing factor. Many students, including Sydney, first attempted to create an equally likely scenario by halving each of the row marginal frequencies, but then either noticed or were directed to the specified column marginal frequencies that were not the sum of these halved values. After the students recognized the restrictions, some tried a guess and check approach, but were not successful. Sydney struggled to begin with and after asking what the problem was asking, he suggested the ratio of 102:188 (marginal relative frequency for the upper shelf) would have to be equal to the RCRFs (upper shelf % of store-brand cereal and upper shelf % of name-brand cereal). He computed this to be 54% (see Figure 20). Then he divided 118 by 100 to get what 1% would be and multiplied that by 54 to get 63.72 and rounded it to 64 for the number of store-brand boxes on the upper shelf. Sydney next determined the 38 by subtracting 64 from 102 and verified his answer by dividing 38 by 70 and getting the same 74. Students created solutions for part (b) and (c) of task 4 in similar ways as they did for more and less likely situations for the previous tasks; however, they recognized the limits imposed by the specified marginal frequencies.

Figure 20

Sydney's Solution for the Cereal Task (IV#4, Task 4(a))

A	B



Note. Panel A: Sydney's calculations for 4(a). Panel B: Sydney's final answer for 4(a).

Because of time limitations, only two participants (Sydney and Jessie) completed part (a) of task 5, which asked for an equally likely situation with a focus on CCRFs (see Figure 21). Nobody completed part (b) or (c) of task 5. For part (a), Sydney applied the same approach he used in part 3(a), using RCRFs (P1FD, WD), and Jessie used his scaling-up approach where he recognized he could divide the marginal frequency for Drug C by 4 to get the marginal frequency of Drug D. Then he used this relationship to divide the given Drug C joint frequency of 32 by 4 to get 8 for the corresponding Drug D joint frequency and also to multiply the derived Drug D joint frequency of 56 by 4 to get 224 for the corresponding Drug C joint frequency.

Figure 21

Solutions Using Participant Strategies (IV#4, Task 5(a))

A	B

Note. Panel A: Sydney's work for 5(a). Panel B: Jessie's work for 5(a).

The question suggested to condition on columns, but both participants focused on rows. This may be due to the previous questions, which suggested to condition on rows. This may be an indication of the difficulty in switching directions rather than recognizing that in an equally likely situation, either direction can be used. When participants were asked about whether they thought an equally likely conclusion based on conditioning in one direction implies the same conclusion for conditioning in another direction, students varied in their responses but collectively were unsure. This type of question could provide be used by teachers as an extension to support differentiation. There is no learning trajectory that describes how students develop an understanding of bivariate data (Casey, Albert, et al., 2018). Understanding that equally likely in one direction implies equally likely in the other direction is something that could very well be a later part of the learning trajectory. Because this property can be visualized with a mosaic plot, it is reasonable to think that further work with mosaic plots may develop these ideas for students.

Summary of Reasoning with Contingency Tables

Students used their quantitative and numerical reasoning abilities, including fractions, ratios, and proportional reasoning, when working with complete contingency tables. The consistent structure of the contingency table and way the questions were asked likely influenced on students garnering their proportional reasoning, and they demonstrated both part-part (odds) and part-whole (risk) approaches. Past studies have used more general questions about variables. For example, Batanero et al. (1996) used the following question for the smoking problem:

> Using the information contained in this table, would you think that, for this sample of people, bronchial disease depends on smoking? Explain your answer.

In contrast, the questions I posed asked students to compare particular categories of variables. For example, for the smoking problem, I would have asked:

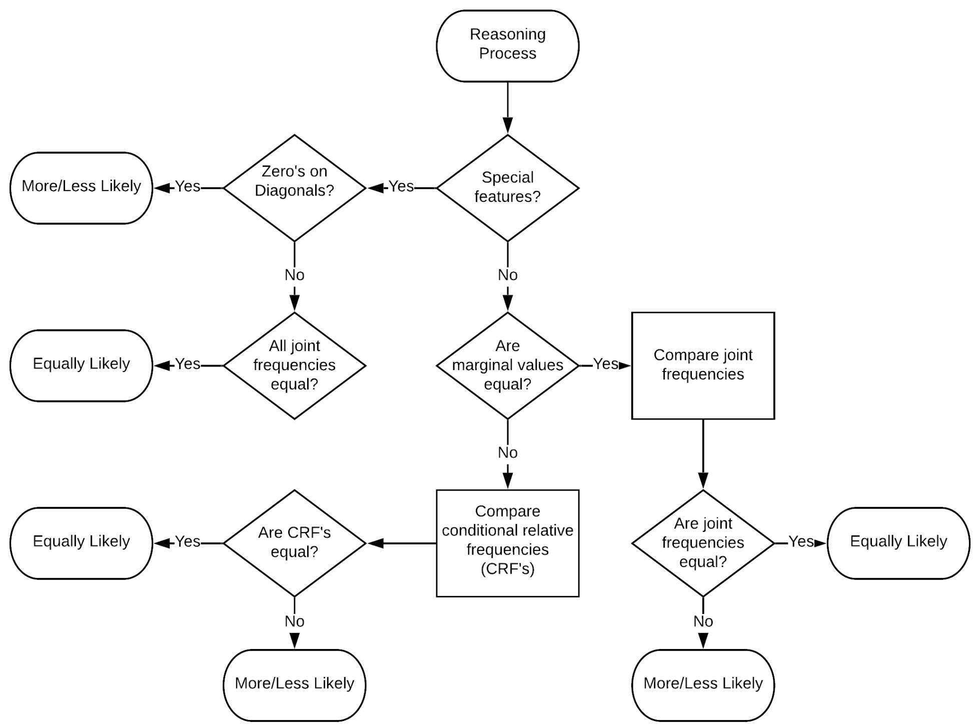
> Considering the information contained in this table, are people who smoke more, less, or equally likely than those who do not smoke to have bronchial disease as opposed to not have bronchial disease?

Because I asked specifically about comparing categories of a variable as opposed to comparing the variables themselves, the students compared conditional values to one another rather than to marginal values.

Students' reasoning seemed to follow a general pattern (see Figure 22) where they adopted a particular strategy and augmented it when more efficient calculations made sense. They first recognized if there were special features of the problem and if so, they used a strategy that simplified calculations and comparisons and often compared part-to-part ratios (P3). If no special features existed, they generally used a part-to-whole approach (P1), dependent on whether the marginal frequencies were equal.

Figure 22

Participant Reasoning with Contingency Tables



When the participants compared the row conditional relative frequencies CRFs, they used a variety of approaches. Table 18 summarizes the participants' approaches.

Table 18

Participant Strategies

Participant	Name	Approaches when Comparing Conditional Relative Frequencies
1	Jamie	Fraction notation converted to a decimal (P1FD)
3	Zander	Division equations resulting in a decimal (P1ED)
4	Sydney	Fraction notation converted to a decimal (P1FD)
6	Hayden	Fraction notation finding common denominators (P1F)
8	Jessie	Scaling rows to equate marginal frequencies (P1SU)

Students often re-read parts of the task or re-stated the question and considered different components of the contingency table as candidates to compute proportions. They experienced some challenges coordinating the constituent components, especially when the wording or structure of the question was changed (e.g., not providing more/less/equally likely, asking for CCRFs vs. RCRFs). Researchers suggested that teachers help students develop flexibility so they can consider conditional relative frequencies in either direction (Watson & Callingham, 2014).

Participants' Reasoning with Mosaic Plots

My second research question aims to understand how students reason about (in)dependence with mosaic plots and contingency tables. Mosaic plots were identified as a useful tool when reasoning with data in contingency tables(Oldford & Cherry, 2006; Pfannkuch & Budgett, 2017). Reasoning with contingency tables and mosaic plots includes both complete and incomplete contingency tables. Interview #3 included complete contingency tables with mosaic plots, and interview #5 included incomplete contingency tables with mosaic plots.

Reasoning with Complete Contingency Tables with Mosaic Plots (IV#3)

The tasks for IV#3 along with answers and goals of the interview are included in Appendix D. The goal of the third interview was to assess participants' ability to create a valid Venn diagram from a contingency table, instruct them on how to create a mosaic plot, and to reveal reasoning about (in)dependence with completed contingency tables when an associated mosaic plot was provided. The first several tasks requested students to create Venn diagrams and initially, successful completion of this portion was planned to be necessary for participants to continue. However, all students struggled to create an accurately labeled Venn diagram, likely because they were accustomed to creating them for two defined sets that have an intersection and using them to sort the elements (e.g., characteristics of birds and bats). This might include using the area inside each circle alone and their intersection to label the three different number of items, but not the area outside of the circles. With a contingency table, there are four joint frequencies that need to be included on the Venn diagram, and the area outside of the circles must be considered. Thus, to reason about (in)dependence from a Venn diagram, numbers from areas of unlike shapes need to be compared. I hypothesize that Venn diagrams may not be well suited as a representation for data of categorical variables summarized in contingency tables to reason about (in)dependence. This coincides with Cherry & Olford (2002), who claimed that Venn diagrams ground abstract set operations and do not ground probability, but a mosaic plot does ground probability. On the other hand, students may be unfamiliar with considering the universal set, and using Venn diagrams in combination with a contingency table might be helpful. The analysis of the participants' work with Venn diagrams is beyond the scope of this study, and after discussions with other researchers, I changed the criteria for continuation to correct construction and creation of a mosaic plot. All participants met these new criteria, which they demonstrated through the first three tasks. The remainder of the interview (tasks 4 to 11) required students to reason about (in)dependence with contingency tables that were complete and included an accompanying mosaic plot.

Tasks with Contingency Tables and Associated Mosaic Plots (IV#3)

Tasks in IV#3 include completed contingency tables similar to Figure 23A, which has an explanatory variable of drug type (A/B) and a response variable of adverse effects

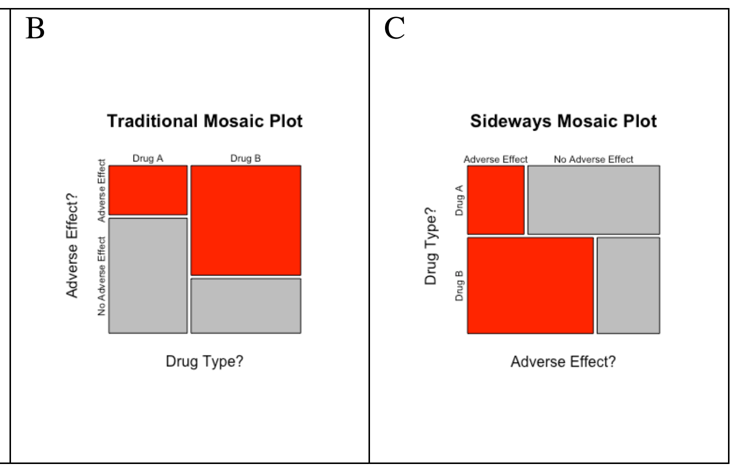
(Y/N). Traditional mosaic plots have a vertical orientation of bars that vary in their segmented heights when there is an association of variables (see Figure 23B). A traditional mosaic plot shows the marginal distribution of the explanatory variable on the horizontal axis. With the table structured with the explanatory variable in rows, the vertical and horizontal components do not align across the contingency table and a traditional mosaic plot. This makes coordination between the representations unnecessarily challenging and induces the Stroop effect, which is when two different representations create cognitive conflict.

Figure 23

Contingency Table and Mosaic Plots with Different Orientations

A

		Adverse effect?		Row
		Yes	No	Totals
Drug Type?	Drug A	60	140	200
	Drug B	187	93	280
Column Totals		247	233	480



Because I used a structure for a contingency table where the explanatory variable categories are the rows, I chose to use what I am calling a sideways mosaic plot (see Figure 23C), which has a horizontal orientation of bars that vary in segmented widths when the variables are not independent. This orientation includes a clear depiction of the marginal distribution of the explanatory variable on the vertical axis. This makes coordination between the representations less challenging, allowing the students to more easily make connections of corresponding parts.

Tasks 4—11 in IV#3 were the same tasks as in IV#2, but they had a mosaic plot for the given contingency table included. Table 13 (p. 85) summarizes the tasks along with some different aspects of the problems.

Applying the Framework to IV#3 Data

Participants answered questions posed in the tasks using only the contingency tables, only the mosaic plots, of a combination of both. Reasoning with only the mosaic plot, without numbers, requires students to reason quantitatively, not numerically because the mosaic plots I included do not have numbers or scales. When participants used the numbers in the contingency tables, they used similar reasoning and strategies to those they used in IV#2. Researchers have found that when students work with numbers and visual information together, mathematics learning is optimized (Park & Brannon, 2013). Table 19 summarizes the codes assigned for each participant across all of the tasks that asked students to reason with contingency tables and accompanying mosaic plots in IV#3.

Table 19

Framework Application for Complete Contingency Tables and Mosaic Plots (IV#3)

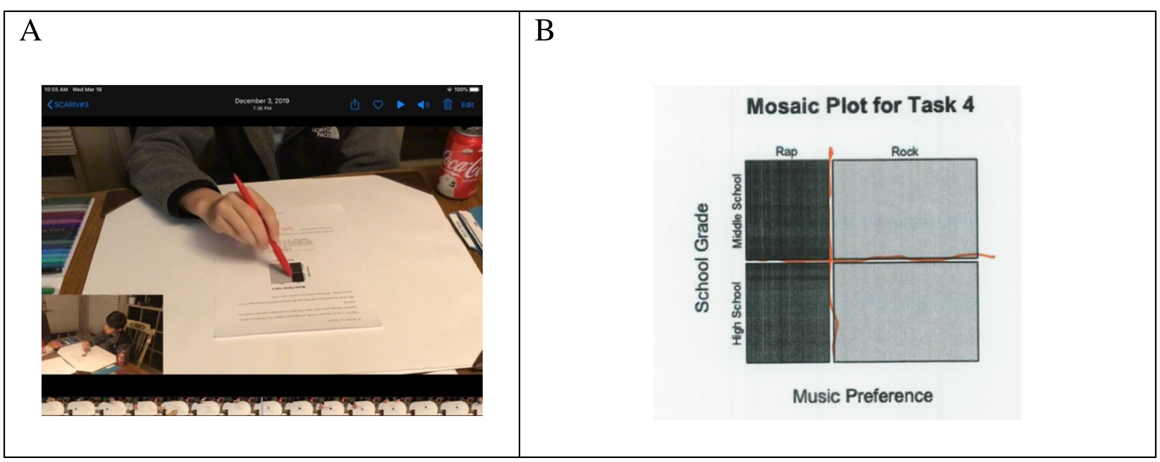
Task	Jamie(P1)	Zander (P3)	Sydney (P4)	Hayden (P6)	Jessie (P8)
4	P1SF, P1MP	P3RMP	P1FMP	P1F, P1SMP	P3R, P1F
5	P1FMP	P3MP, P3FD	P1FMP	P1F, P1MP	P3MP, P1FMP
6	P1MP	P3MP, P1SW, P1SMP	P1MP, P1SW	P1SF, P1SMP	P1MP
7	P1MP, P1FD	P1MP, P1ED	P1FMP, P1FD, P3MP	P1BF	P1MP
8	P1MP, P1FD	omitted	P1SMP	P1SF, P1SMP	P1MP

Task	Jamie(P1)	Zander (P3)	Sydney (P4)	Hayden (P6)	Jessie (P8)
9	P1MP, P1FD	omitted	P1FMP	P1F	P1MP
10	P1MP	omitted	P1SUMP	P1F, P1MP	P1MP
11	P1FMP	omitted	P1FMP	P1F	P1MP

Task 4 had equal proportions, and participants initially either used the contingency table alone (Jamie, Hayden, and Jessie) or the contingency table in conjunction with the mosaic plot (Zander and Sydney) to conclude independence. Jamie first used only the numbers in the contingency table, and then when he used the mosaic plot, he traced the vertical line that cuts through both rows of the mosaic plot, identifying it as a feature that signified equally likely. Similarly, Hayden identified the vertical line, but he also identified the marginal frequencies were the same because the horizontal line went through the middle (see Figure 24).

Figure 24

Work with a Vertical Line on a Mosaic Plot (IV#3, Task 4)



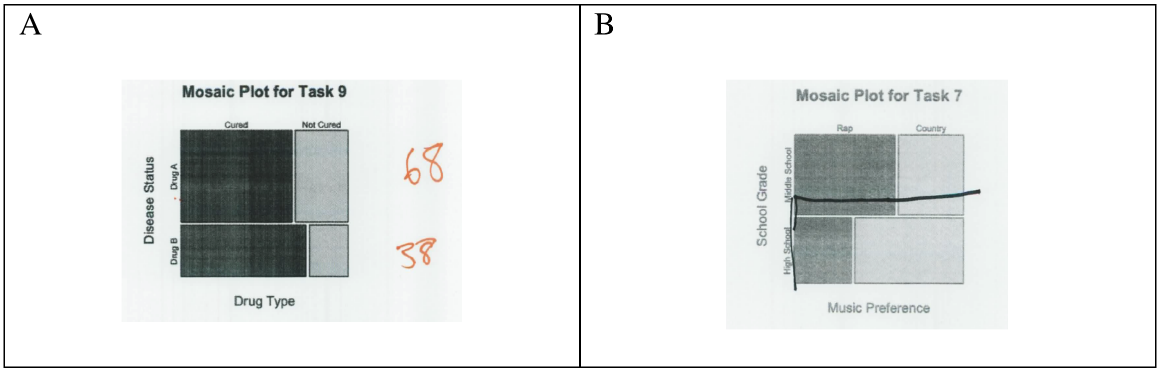
Note. Panel A: Jamie traces the vertical line. Panel B: Hayden marked vertical and horizontal lines.

The vertical line that goes through both rows of the mosaic plot is something that all students recognized as a feature of independence; however, Sydney and Hayden were not sure they could use this criterion when the marginal frequencies, or the heights of the rows, were different. When a student reasons with a new representation, connections have to be made with other representations, and this challenging cognitive process can create obstacles (Dreher & Kuntze, 2015). Obstacles are necessary for learning and determining when the vertical line criteria can be applied seems to be an area that is challenging for some students. Questions related to this should be included in instruction because encouraging students to actively create connections between multiple representations advances learning (Dreher & Kuntze, 2015).

For example, with a later task where the marginal frequencies were different (task 9), Sydney identified that because the horizontal line was not down the middle, he would have to calculate the numbers, and Hayden said he did not use the mosaic plot because the total number of observations for the rows was different (see Figure 25A). On the other hand, when working with another task where the marginal frequencies were different (task 7), Jessie mentioned that it did not matter if the row totals were different or the same, drawing a horizontal line across the middle of the mosaic plot (see Figure 25B). Like Jessie, both Jamie and Zander recognized the mosaic plot could be used whether the marginal frequencies were the same or different.

Figure 25

Work with Mosaic Plots Recognizing Marginal Frequencies



Note. Panel A: Hayden noticed different marginal frequencies. Panel B: Jessie noted equal marginal frequencies.

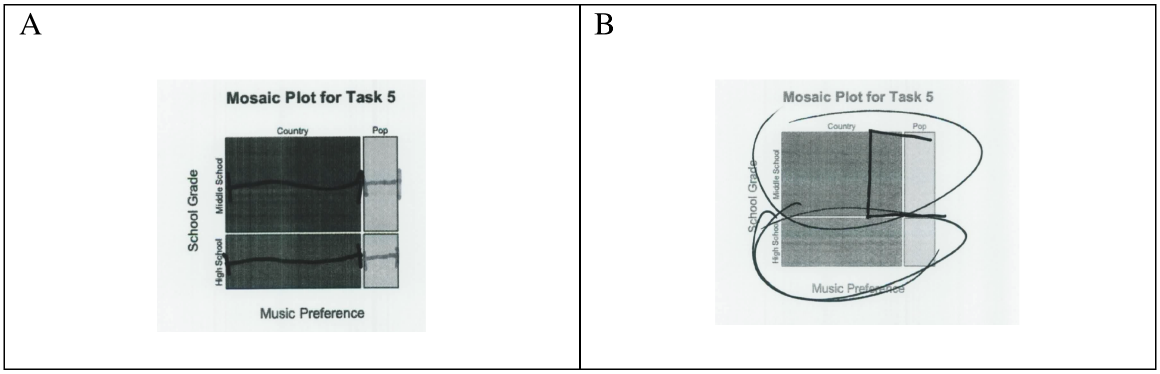
Sydney and Hayden did not demonstrate the same depth of understanding of the usefulness of a mosaic plot. It is interesting to note Sydney and Hayden's preference for working with the numbers in the contingency tables as opposed to using drawings. Sydney mentioned he first looked at the mosaic plot in task 4 and saw the RCRFs appeared equal, but he recognized that this appearance does not mean they are equally likely because "They could be one apart." At the end of the interviews, he admitted that he does not like drawing things or using pictures and that if problems do not have numbers, he does not like them. At the beginning of task 4, Hayden momentarily considered using the mosaic plot, pointing to it and saying "from, I guess", indicating he was going to use it to reason with, but then he fairly quickly changed his plan and said, "Actually, I'm going to look at the contingency table." At the end of the interview, Hayden said the mosaic plot provided an estimate, and using the numbers in a contingency table he provided an exact answer. Furthermore, he said, "For math you shouldn't really estimate" and that for a test, he was not going to look at a picture.

It is reasonable to think that Sydney and Hayden's preferences for working with numbers rather than pictures contributed to their more limited understanding of a mosaic plot where they thought the marginal frequencies needed to be equal in order to use the mosaic plot to determine (in)dependence. Sydney and Hayden's affects (attitudes, interests, and values) seemed to contribute to lessening their engagement with visual information, thus impacting their depth of understanding of mosaic plots. There could be other reasons why they thought equal marginal frequencies were a requirement for using the mosaic plot. One possibility is that equality of marginal frequencies was a significant part of student reasoning in the previous interview (IV#2), as they used this criterion to compare whole number joint frequencies. The participants may have been transferring this reasoning to working with mosaic plots.

Beginning with task 5, Jamie used the mosaic plot in conjunction with the numbers in the contingency table. He first identified the fractions for the RCRFs, but then rather than calculating the decimal values, he realized he could use the mosaic plot. Jamie recognized the vertical line that went through both rows signified the percentages were the same (P1FMP). Initially he thought he could only use the mosaic plot to determine (in)dependence when the situation was one that was equally likely, but he soon realized he could use it to recognize more and less likely situations, too. Zander and Jessie started reasoning with only the mosaic plot in task 5. They both began by comparing the parts of the mosaic plot within and across the rows, considering the ratio of the parts of the top row in comparison with the ratio of the parts of the bottom row. Zander noted he was comparing the horizontal distances of each of the blocks of the mosaic plot (see Figure 26A).

Figure 26

Participant Solutions Using Horizontal Distance (IV#3, Task 5)



Note. Panel A: Zander's P3MP work. Panel B: Jessie's P3MP work.

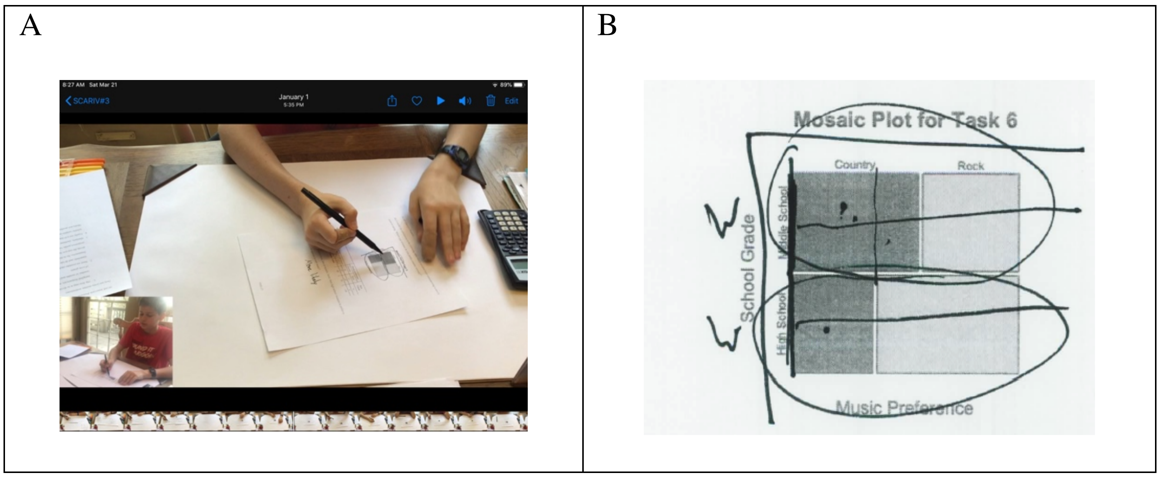
Jessie similarly compared the parts of the mosaic plot and furthermore provided an example of how the mosaic plot would look if the situation were less likely instead of equally likely (see Figure 26B). Both Zander and Jessie used the linear distance of the parts, not the area of the parts, to compare them in task 5.

Sydney and Hayden, on the other hand, often used the areas of the blocks of the mosaic plot for comparison. This could be another factor that contributed to their thinking the marginal frequencies had to be equal in order to use the mosaic plot. The ratio of the areas can always be compared to determine (in)dependence, but the areas themselves can only be compared when the marginal frequencies are equal. Working with mosaic plots affords clear connections with geometry and measurement where the difference of measuring the one-dimensional linear distance and two-dimensional areas have implications.

For task 6, Jessie said he did not have to look at the information and after reading the question he used the mosaic plot to explain it is more likely. He justified his answer saying, "This bar is greater than this bar," pointing to the middle school/country and high school/country portions of the mosaic plot. Jessie compared the horizontal lengths of the parts for country and when asked to further explain, Jessie turned the mosaic plot 90 degrees (see Figure 27A) and drew an axis to display the mosaic plot as a bar chart with a traditional orientation with vertical bars to explain the heights of the bars were different.

Figure 27

Jessie's Work with Mosaic Plot Dimensions (IV#3, Task 6)



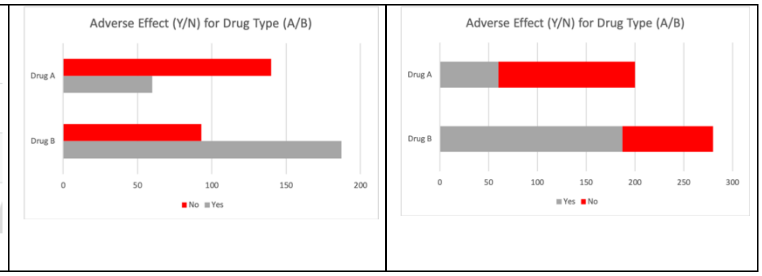
Note. Panel A: Rotating paper 90 degrees. Panel B: Mosaic plot with the Cartesian plane drawn and "W" for width.

The width (or horizontal length) is an unusual metric to attend to for a bar chart, and Jessie struggled with what to call the height of the rows of the mosaic plot, initially choosing width and labeling it with a "W" and later calling it a length and height (see Figure 27B). One challenge with a sideways mosaic plot may be overcoming normative graphical displays using the Cartesian plane where the bars are vertical rather than horizontal. This is a similar challenge that exists when considering segmented or side-by-side bar charts with a different orientation (see Figure 28). However, a benefit of this orientation is the numerical values are in the same orientation as a traditional number line. It is unclear whether aligning the orientation of graphs to coincide with the contingency tables may be helpful. Research that looked at bar graphs and mosaic plots with contingency tables included only traditional graphs with vertical bars (Casey, Albert, et al., 2018; Casey, Hudson, et al., 2018).

Figure 28

Contingency Table and Bar Graphs with Horizontal Orientation

		Adverse effect?		Row
		Yes	No	Totals
Drug Type?	Drug A	60	140	200
	Drug B	187	93	280
Column Totals		247	233	480

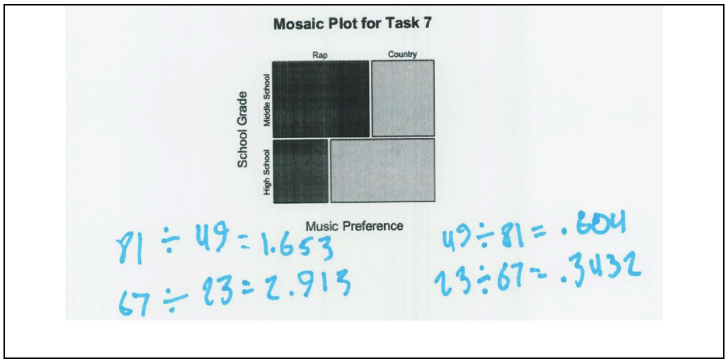


For task 6, Zander used the area instead of the linear distance he had used in task 5. He first used the mosaic plot, pointing to the middle school/country block (P3MP), and then noticed the marginal frequencies were the same, so he compared whole numbers (P1SW). When further explaining how he used the mosaic plot, he indicated he was using an odds type of reasoning by noting that the middle school/county part is bigger than the middle school/rock, tracing the perimeter of the middle school parts, and then he mentioned you could also compare the areas of the country column between the middle school and high school categories (P1SMP). This reasoning demonstrated that Zander understood that the area of the parts of the mosaic plot can be used in two different ways. First, the ratio of areas can be compared, which is similar to an odds approach and is always appropriate. Secondly, the areas of two parts for one column can be compared when the RMFs are equal or the horizontal line is located in the middle, which is the same as comparing the joint frequencies for one category of a variable (country) when the marginal frequencies for the categories of the other variable (middle school/high school) are equal. Working with mosaic plots may provide an opportunity for students to make connections to geometry and measurement, which is an area where many students have difficulties. Students often have a difficult time and conflate the measurements of length and area. For example, students have a difficult time differentiating between area and perimeter (Curry & Outhred, 2005; Lehrer, 2012). Work with mosaic plots allows students to consider different measurements, and this can be an opportunity to refine their mathematical language (e.g., area, horizontal, vertical, length, height, width). Understanding key terms and using them appropriately is important for success in more advanced mathematics. Working with mosaic plots may also develop flexibility with different visual representations beyond graphs with a traditional Cartesian plane.

Task 7 was the last task Zander completed in IV#3, and he used an unconventional whole-to-part ratio for his numerical solution. He first solved it quantitatively, using only the mosaic plot. Zander used the mosaic plot to determine more likely and then verified this with equations using numbers from the table (see Figure 29). He first computed the marginal frequencies divided by the joint frequencies, compared them numerically, but struggled to interpret them in context. Then he computed proportions or percentages and was able to explain them in context. Here is an example where part-to-whole ratios appear easier to explain than another type, in this case as a whole-to-part ratio.

Figure 29

Zander's Work with Equations (IV#3, Task 7)



For the remaining tasks, the participants continued to use the strategies they developed in IV#2 along with the mosaic plot, where some students preferred to use the mosaic plot alone. Hayden focused first on the contingency table, using his strategy with equivalent fractions and Sydney used the contingency table in conjunction with the mosaic plot. Jamie and Jessie relied on solely the mosaic plots, finding them both useful and efficient when determining (in)dependence. At one point Jessie said "Man, it's really easy when you have a mosaic plot" and "It seems like the best way to solve the problem." He explained they are already evened out, do not require multiplication, and can be directly compared.

Reasoning with Incomplete Contingency Tables with Mosaic Plots (IV#5)

Appendix F identifies the goals of this interview along with tasks and answers. There were three goals of the fifth interview (IV#5), which like the fourth interview included contingency tables with incomplete information. First, IV#5 assessed participants' abilities to draw representations from the information in the contingency table, including a representation of their choosing as well as a mosaic plot. Next, IV#5 aimed to gain awareness of the participant's current understanding of some terms (association and independence, category, and variable). Finally, the main goal of IV#5 was to understand how students reasoned about (in)dependence with contingency tables that were incomplete when a mosaic plot was provided and either necessary or useful to solve the problem. IV#5 included four tasks with the same context as some of the tasks in past interviews.

Tasks for the Fifth Interview (IV#5)

Tasks in the fifth interview requested students to draw representations, discuss their understanding of key terms, and reason about incomplete contingency tables in conjunction with a mosaic plot. The first task in IV#5 used the same context as task 1 in IV#4, a survey of adults and children asking whether they preferred dogs or cats; however, it included a complete contingency table with a more likely situation and requested participants to create a representation of their choosing and then construct a mosaic plot. Later parts of this task had students create a mosaic plot for less likely and equally likely situations using the same context. Next, in task 2, students were asked for their understanding of some terms (association, independent, category, variable).

The remaining tasks were based on tasks in IV#4, and each used the same contingency table but also included a corresponding mosaic plot. The questions asked the students to complete the missing entries in the table to represent different situations of less likely and equally likely. Tasks 3(a) and (b) in IV#5 are the same as tasks 2(a) and (b) in IV#4, but they have a mosaic plot included (see Table 20). Task 4 in IV#5 includes the same contingency table as task 4(a) in IV#4 that students struggled to complete.

Table 20

Task Summary for Incomplete Contingency Tables and Mosaic Plots (IV#5)

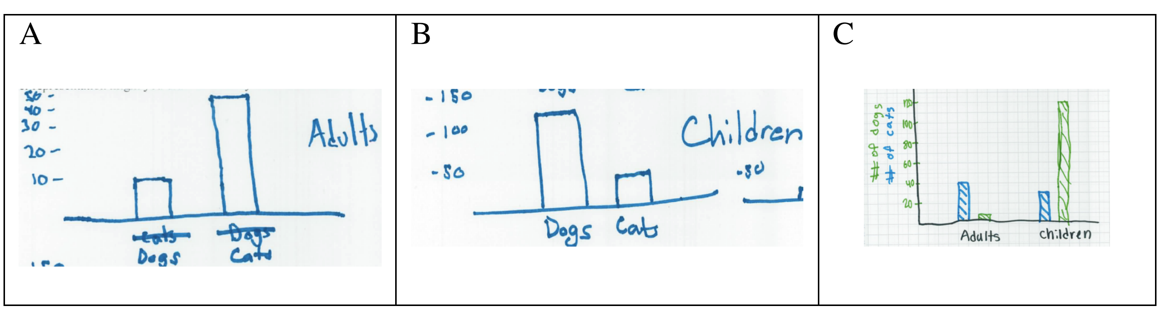
Task	Context Description	Completed Cells	Direction and Base Row/Column	Row Marginal Frequencies	Order (a)/(b)
3	Exercise Frequently/Seldom Sleeping Problems/No Problems	Row Marginal Frequencies	RCRF, 1, 2	Different	Less/Equal
4	Cereal Brand Name/Store Shelf Upper/Lower	All Marginal Frequencies	RCRF, 1, 2	Different	Equal

Creating Representations - Graphs and Mosaic Plots (IV#5, Task 1)

When given the option to draw any representation of their choosing to represent the situation in task 1, participants either drew side-by-side bar charts using frequencies (Jamie, Zander, and Jessie), or they drew a mosaic plot (Sydney and Hayden). Jamie noticed the different marginal frequencies and drew separate graphs for children and adults (see Figure 30A and Figure 30B), saying that the different scales made it easier to look at because if he used the same scale, the graph for the adults would be really big or the graph for the children would be really small (see Figure 30C).

Figure 30

Participants' Side by Side Bar Graphs



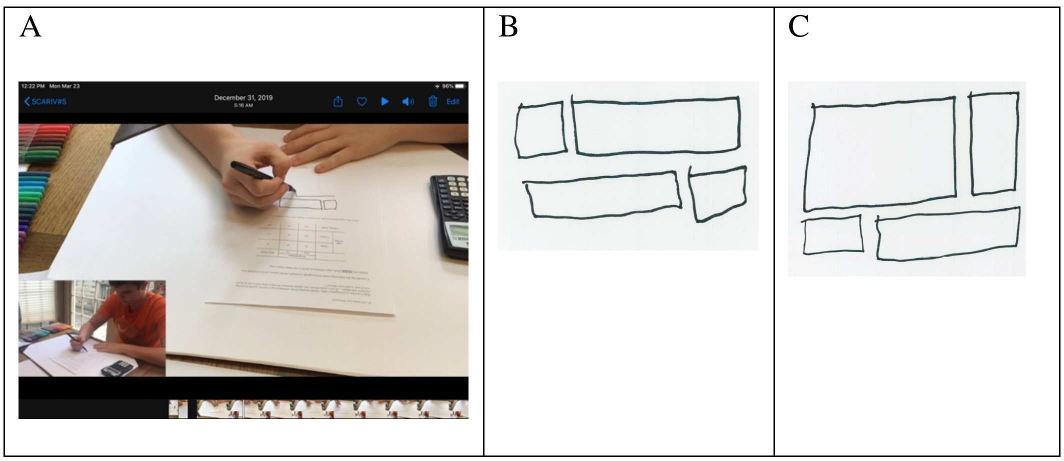
Notes. Panel A: Jamie's graphs for adults. Panel B: Jamie's graphs for children. Panel C: Zander's graph.

The bar graphs the students selected to draw are consistent with some findings in the literature. Work with middle grades students (Casey, Hudson, et al., 2018) and secondary teachers (Casey, Albert, et al., 2018) revealed that side-by-side bar graphs were more commonly used than segmented-bar graphs and that scales using frequencies rather than relative frequencies were more common. Researchers suggested it is correct to use relative frequencies (proportions or percentages) as scales for a bar graph when reasoning about association of categorical data (Casey, Albert, et al., 2018). In essence, this is what Jamie was trying to do with using different scales for his graphs. Adjusting the scale of the graph to percentages makes the bar graph more similar to a mosaic plot where a relative comparison of the proportions is supported. A mosaic plot provides additional information beyond a percentage segmented bar graph, displaying the marginal distribution through adjusting the width of the bars.

Sydney and Hayden chose to draw a mosaic plot in response to the first task, likely because they had experience creating them in the previous interview. Sydney drew separate boxes or tiles for the parts of the mosaic plot without using a grid, keeping the heights of the blocks the same for each row, despite unequal numbers of adults and children. This aligns with his preference for equal marginal frequencies and his thoughts that they have to be equal to use the mosaic plot. He started using areas in relation to frequencies, first drawing the adult/dog box as a square (10) and then iterating the square four times to draw a rectangle for the adult/cat portion (40). Next, he transitioned from a focus on area to horizontal length where he used the proportion of 4/5 to determine the width of the child/dog portion (see Figure 31A). This representation was similar to a percentage bar graph with a horizontal orientation. After completing the mosaic plot for task 1(a), Sydney recognized the marginal frequencies were different and his drawing did not reflect this difference (see Figure 31B). In the subsequent portions of the tasks, the mosaic plot that Sydney drew accounted for different marginal frequencies (see Figure 31C).

Figure 31

Sydney's Work Using Proportions to Draw a Mosaic Plot (IV#5, Task 1)



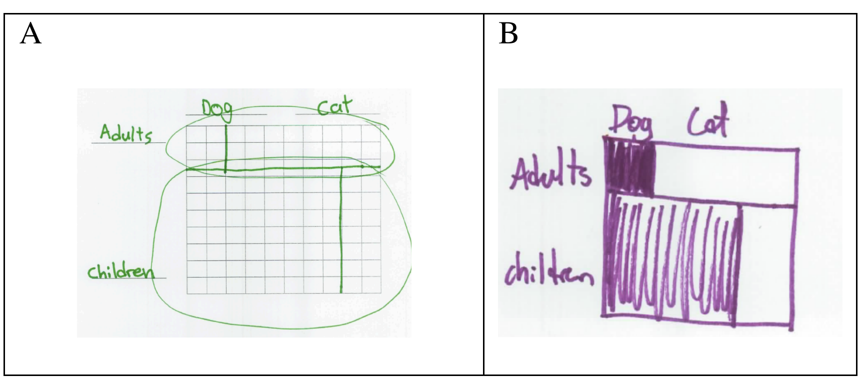
Note. Panel A: Using 4/5 to determine the width. Panel B: Mosaic plot for part (a). Panel C: Mosaic plot for part (b).

Marginal frequencies seem to hold a place of importance for students; yet, they struggle with knowing when these values matter and how to use the fact they are the same or different. This information could inform the sequencing of tasks in order to make clear connections.

Hayden initially attempted to draw a Venn diagram but struggled to label the circles in a way that there was a particular number of observations in the overlapping area. His struggle is more evidence for the difficulties students have when using Venn diagrams with contingency tables. Students should not be expected to reason with Venn diagrams exploring the association of two categorical variables without some explicit instruction. Having them create a Venn diagram with some guidance might aid their understanding, but there is not enough research in this area to know the nuances of how this may or may not support learning. After being unsuccessful in creating a Venn diagram, Hayden changed plans and decided to draw a mosaic plot. He used a grid, labeled the categories of each variable, drew a horizontal line to separate adults and children, and then used a proportion (4/5 and 1/5) to partition the adults and children into those who prefer dogs and those who prefer cats (see Figure 32A). Hayden referenced both a one-dimensional measurement of length and a two-dimensional measurement of length as he was creating the mosaic plot. When asked to summarize his reasoning, Hayden recognized the linear distance as the most important metric. Hayden's actions suggest that students should not be expected to reason with drawn representations after creating or seeing them for the first time. Rather, continuing to draw representations across multiple tasks can support and extend learning.

Figure 32

Participants' Work Creating Mosaic Plots (IV#5, Task 1(a))



Notes. Panel A: Hayden's mosaic plot. Panel B: Jessie's mosaic plot.

All other participants created mosaic plots without assistance after they created other representations. Similar to Hayden, Jamie and Zander used a grid to create the first mosaic plot. Jamie used the linear horizontal distance to determine where to draw the vertical lines. On the other hand, Zander, who had reasoned flexibly with distance and area in a previous interview, recognized that each square in the grid represented two observations, and he focused on the area to determine where to partition between dogs and cats. Jessie, who had drawn a mosaic plot in IV#4 to reason with an incomplete contingency table, drew a mosaic plot without a grid (see Figure 32B).

Understanding of Terms

Participants generally thought of the word *independent* as "alone" or "by itself" and *association* as "together" or "in a group." Some students also mentioned independent variables as ones that change other numbers but are not changed by other numbers. When asked to consider whether more or less likely situations and mosaic plots were independent or associated, most students said independent because the proportions are not the same. Zander, on the other hand, chose associated for all situations (more likely, less likely, and equally likely) because the numbers and the parts of the mosaic plot could be connected, so he saw them as going together. For equally likely situations, other participants selected association because the proportions were the same. Students' understanding of the terms *association* and *independent* may be a challenging factor in developing an understanding of what these terms mean in a statistical sense. Possibly a focus on the relationship of the variables and how they might work together when associated could be helpful.

Participants' understanding of categories and variables seemed to be less consistent. Participants thought of a category as either the variables (e.g., age group, pet preference), the categories of each variable (dog, cat; adult, child) or the combinations of categories for each variable (adult/dog, adult/cat, child/dog, child/cat). A variable was thought of as a number represented by *x* or *y* that can change and can be independent or dependent, or more simply the unknown answer to the problem. Most of the participants identified variables as the numbers in the contingency tables; however, Zander and Jessie had other thoughts. Zander recognized the independent variable as adults and children, and Jessie said the variable was the "unknown." Jessie gave an example of the "unknown" as the missing numbers in an incomplete contingency table or the entire mosaic plot when the problem provided a contingency table and asked him to create a mosaic plot. In summary, participants could generally identify the categories, but they struggled to clearly identify the two variables in the problem as age group and pet preference.

This gap signifies a need for teachers to help students understand that a contingency table represents bivariate data using summary counts that combines information for two variables. It may be helpful to connect students' algebraic understanding of independent and dependent variables to explanatory and response variables with a caution that these are not always clearly included in contingency tables and questions about design should be addressed. For example, a contingency table with middle and high school students as categories of one variable and country and rap as categories of another variable could have resulted from a survey of middle and high school students asking their preference between country and rap music. However, the same information could have resulted from a survey taken at a country and rap concert asking whether people were in middle or high school.

There are additional challenging terms when working with contingency tables (e.g., conditional, relative, marginal, frequency, proportion). The understanding of statistical language is beyond the scope of this study; however, educators should be cognizant of the difficulties students may have and take time to assess students' current understanding of terms so they can help students develop the understanding that is needed when working with data from categorical variables and represented with counts in contingency tables.

Applying the Framework to IV#5 Data

There were two tasks that included mosaic plots and incomplete contingency tables. Task 3 used the exercise and sleep problem from IV#4 and required students to use the mosaic plot to estimate less likely and equally likely situations. Task 4 used the challenging cereal and shelf problem from IV#4, and the mosaic plot was not necessary but helpful. Table 21 summarizes the codes assigned for each participant across all of the tasks in IV#5.

Table 21

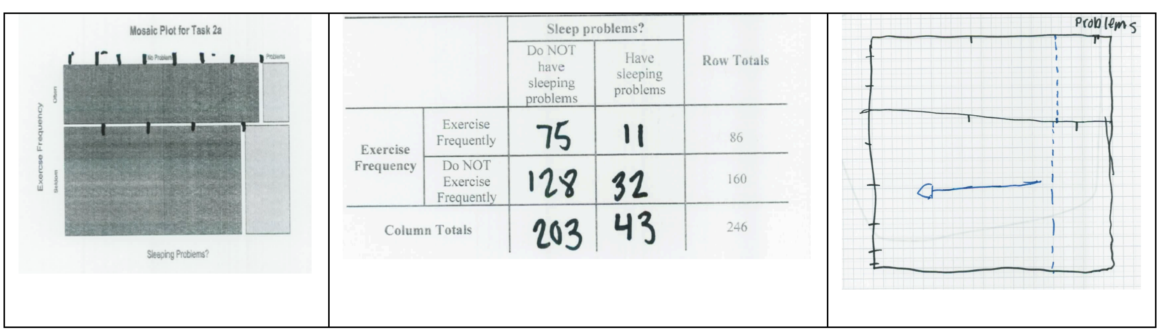
Framework Application for Incomplete Contingency Tables with Mosaic Plots (IV#5)

Task	Jamie(P1)	Zander (P3)	Sydney (P4)	Hayden (P6)	Jessie (P8)
3(a)	P3AN, P1FMP	P1FMP	P1BF, P1FMP	P1FMP	P3AN
3(b)	P1BFMP	P1BFMP	P1BFMP	P1BFMP	P1BFMP
4(a)	P1MPTE	*P2MP, P2MPD*	P2MPD	*P1MPTE, P2MPD*	StartP1F

For task 3(a), Jamie and Sydney went straight to the contingency table and did not consider the mosaic plot with their first solution as the question requested. Jamie used an all or none approach (P3AN), and Sydney used a benchmark fraction of ½ (P1BF). Providing a mosaic plot with an incomplete contingency further restricts the possible combinations of numbers and requires students to use the mosaic plot to compare areas. When considering the mosaic plot for task 3(a), all students estimated the proportions for each row in the mosaic plot to determine how to split the row marginal frequencies for adults and children (P1FMP). For example, Figure 33A shows Zander's work where he recognized the "have sleeping problems" part for the top row "often" was one of about 8 equal size pieces. Likewise, for the bottom row "seldom," the "have sleeping problems" part was one of about 5 equal size pieces. To estimate the joint frequencies, he divided the marginal frequency by the number of pieces (86 ÷ 8 = 11 and 160 ÷ 5 = 32) (see Figure 33B for the completed contingency table). Zander drew the representation in Figure 33C to show the range of possibilities for the less likely situation when fixing the proportions for those who exercise frequently in the top row. The value for the "do not have sleeping problems" proportion of people who do not exercise frequently just needs to be less than 140 (7/8 of 160).

Figure 33

Zander's Work Using a Mosaic Plot to Determine Frequencies (IV#5, Task 3(a))



		Sleep problems?		
		Do NOT have sleeping problems	Have sleeping problems	Row Totals
Exercise Frequency	Exercise Frequently	75	11	86
	Do NOT Exercise Frequently	128	32	160
Column Totals		203	43	246

Notes. Panel A: Mosaic plot with partitions. Panel B: Completed contingency table. Panel C: Drawn mosaic plot showing a range of numbers.

Zander and the other participants were able to use the mosaic plot to estimate the RCRFs. Compared to creating a mosaic plot from a contingency table, completing this problem demonstrates reversibility and a deeper understanding of reasoning across representations. Part (b) of task 3 provided a contingency table for the situation that was equally likely. All of the participants estimated that the vertical line going through both rows was down the middle and halved each of the marginal frequencies. They recognized it did not have to be down the middle to be equally likely; it just had to be a vertical line signifying CCRFs were the same.

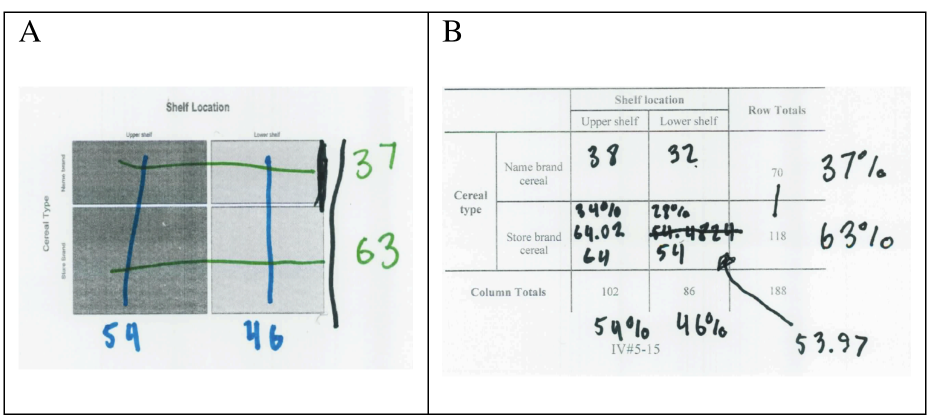
Task 4 specified all marginal frequencies and a situation of equally likely. Other than a trial and error approach, this task required using both row and column marginal frequencies to generate a solution. In the previous interview, Sydney was the only one who solved this problem, and he used one of the column MRFs (54%) and a row marginal frequency (118) to determine a joint frequency (54 x 1.18 = 64). This time, he was not able to determine a solution and asked for assistance. I had him identify where the marginal frequencies appeared in the mosaic plot and reminded him of the strategy he had been using where he found percentages. He then used a similar approach, using a marginal frequency and an MRF, to find one of the joint frequencies.

In this interview, all participants except Zander started with a trial and error approach for task 4. Jamie was able to use trial and error to find a solution, whereas the other participants struggled to find a solution using this approach. After I encouraged them to consider the mosaic plot and marginal values, all participants except Jessie, who seemed fatigued at this point in the problem and asked to stop, were able to find a solution, using a similar approach to what Sydney did.

Zander was the only participant who did not begin by using a trial and error approach. He found the mosaic plot especially useful in this problem (see Figure 34A), and he observed it to "try to think of a way to find numbers that would work" when given all the marginal frequencies. Figure 34B shows Zander's work with the numbers in the contingency table. He calculated all of the marginal relative frequencies (e.g., 70 ÷ 188 = 37%). Then he multiplied marginal relative frequencies to determine the joint relative frequencies for the store-brand cereal (46% * 63% = 28%). He said he multiplied them because he was creating a combined percentage and he knew that dividing or adding could result in a percentage greater than 1. Then he took the joint relative frequency and multiplied it by the total number of cereal boxes to determine the joint frequency (28% * 188 = 54).

Figure 34

Zander's Work Using a Mosaic Plot to Solve the Cereal Task (IV#5, Task 3(a))

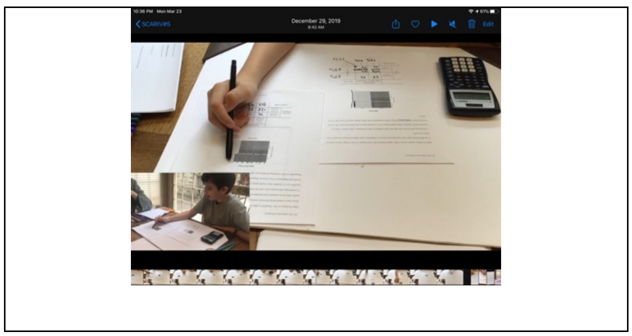


Notes. Panel A: Mosaic plot. Panel B: Contingency table.

When asked how he figured it out, Zander said, "The mosaic plot gave me the idea first," explaining he was trying to find out what percentage the name-brand cereal height was out of the total height (drawing vertical lines on the right of the mosaic plot). He referenced work with 3(a) to explain where the idea originated. For 3(a), looking at the parts in the top row, he estimated that one piece would fit into the other 7 times (see Figure 35). Likewise, looking at the parts in the right column for the current problem, he noticed one piece fit into the other about 2 times (see Figure 34A). Because he wanted to find the exact number, he used the marginal frequencies.

Figure 35

Zander's Work Referencing IV#5, Task 3(a)



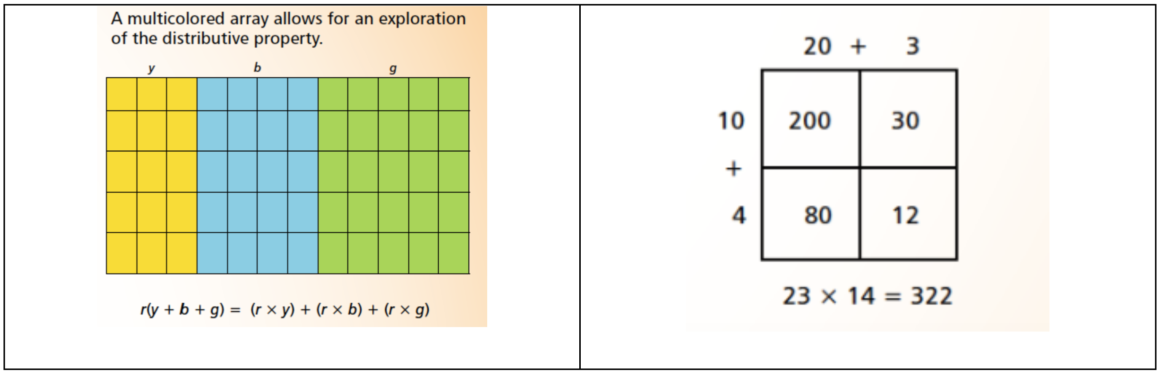
Zander visibly used the mosaic plot to support his reasoning and solved a problem he was not able to solve with a contingency table alone. When asked if he could use just one of the marginal relative frequencies to determine one of the joint frequencies, Zander did not think it was possible. He recognized the 63% as the bottom portion of the entire mosaic plot and initially did not think it had any relevance for just the lower shelf portion of the mosaic plot. At this point, Zander did not recognize that the distributive property was relevant.

The distributive property is not an algorithm (Benson et al., 2013), and unlike how it is presented in China, the underlying principle is rarely made explicit (Ding & Li, 2010) in instruction in the US. Most US high school students come away with the thoughts that the distributive property is about computation and parentheses in expressions. The distributive property, unlike the associative property, includes two different operations. Whereas multiplication and addition are each an operation where the associative and commutative properties hold, the distributive property holds for multiplication over addition (e.g., $a * (x + y) = a * x + a * y$), but not for addition over multiplication (e.g., $a + (x * y) \neq (a + x) * (a + y)$). In other words, the distributive property equates multiplying a sum with multiplying each of the addends separately and adding the results. For example, if you have two parts of something that measure *x* and *y*, multiplying the parts individually by a multiplier *a* and then adding them together (e.g., $a * x + a * y$) will yield the same answer as adding them together first and then multiplying them by the same multiplier *a* (e.g., $a * (x + y)$). Thus, the distributive property is about counting things in different ways rather than just distributing the multiplier to the addends to remove the parentheses.

Researchers found that the distributive property, when viewed numerically and geometrically, is naturally logical to even elementary students (Benson et al., 2013). Using an array, and later a more abstract box model, allowed for a transition from concrete to abstract (see Figure 36). Students were able to "demonstrate their understanding of conservation of area and to discover a geometric interpretation of the distributive property" (Benson et al., 2013, p. 498). Chinese texts include problems across grades that assist students to learn and apply the distributive property (Ding & Li, 2010). For example, in addition to the traditional equation for perimeter, $p = 2l + 2w$, Chinese textbooks include an additional equation, $p = 2\ (l + w)$. and ask students to make connections between the two equations. Some activities related to the geometric interpretation of the distributive property in conjunction to work with contingency tables and mosaic plots might be beneficial to US students.

Figure 36

Array and Box Model

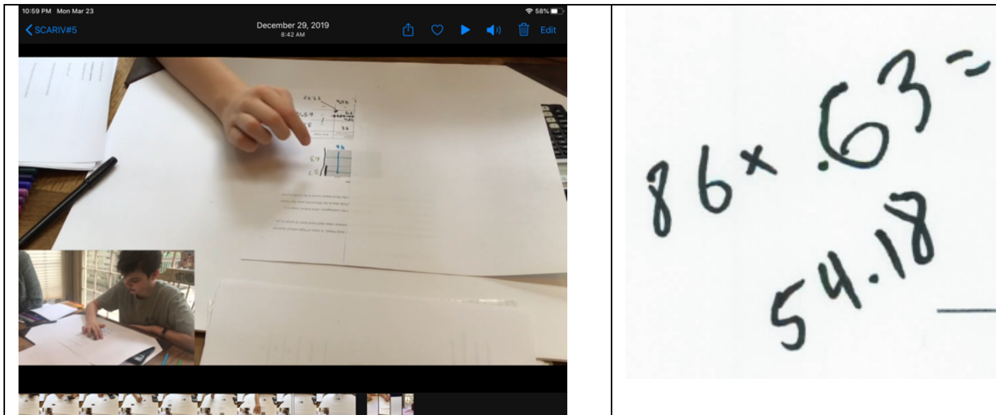


Note: Reprinted from "The Distributive Property in Grade 3?," by C. C. Benson, J. J. Wall, and C. Malm, 2013, *Teaching Children Mathematics, Vol. 19* (8), pp. 499, 501 (DOI: 10.5951/teacchilmath.19.8.0498).

Zander did not recognize the distributive property in the contingency table or mosaic plot. I covered up the portion of the mosaic plot for the upper shelf, the first column. Once Zander realized the same MRFs held for the lower shelf, he used the store-brand cereal MRF of 63% to determine the number of name-brand boxes on the lower shelf (see Figure 37A). Thus, Zander recognized the distributive property in this situation where using the MRF and taking this percentage of the marginal frequency would result in the joint frequency. He recognized this as a more efficient strategy in comparison to multiplying the MRFs and then taking that percentage of the total number of observations (see Figure 37B).

Figure 37

Zander's Work with the Distributive Property with a Mosaic Plot (IV#5, Task 3(a))



Summary of Reasoning with Mosaic Plots

Students were able to create mosaic plots for associated contingency tables according to instruction and subsequently use them to reason about (in)dependence. Creating a representation prior to reasoning with it affords a better understanding where in this instance, the parts of the mosaic plot were connected to the numbers in the contingency table and students began to see how the parts related to one another. In a later interview, all participants were able to create a mosaic plot with no assistance, and in some cases this action enhanced understanding. Having students create mosaic plots deepened their understanding and ability to reason with them.

When using mosaic plots, all the participants recognized that an equally likely situation of independence existed when the two vertical lines that split each row were aligned in the same place where the distance from the line to the edge was the same. However, some students were unclear about relevance for the horizontal line.

When a mosaic plot does not include numbers, either on the axes or within the parts, it invites students to reason quantitatively. The mosaic plots for the tasks in this study did not include any numbers. Work with doctors and patients revealed that accompanying natural frequencies with a visual aid improved students' reasoning (Garcia-Retamero & Hoffrage, 2013). Participants mentioned that it would be helpful for the mosaic plot to include numbers, such as scales, and these could be included as either frequencies or relative frequencies.

Students' reasoning with mosaic plots seemed to follow a general pattern where they looked for a vertical line and used that to conclude equally likely for complete contingency tables or compute joint frequencies using the same proportions for an incomplete contingency table. If there was no single vertical line, they looked at either areas or lengths to determine whether the situation was more or less likely for the category in question. If the marginal frequencies were equal, they could compare the areas or lengths and if the marginal frequencies were not equal, they could compare the linear distances or the ratios of areas within each category.

Participants' Challenges with Wording and Direction

Asymmetric Variables in Task 10

Participants continued to struggle with the wording of questions, especially Sydney who requested to skip task 2(a) in IV#4. This task considers people who do and do not exercise frequently and people who do not and do have sleeping problems, both binary categorical variables. Sydney expressed hesitation with this problem saying, "I don't know about this one; it's worded weird." He indicated he thought he might be comparing teachers and students, which are identified as the people who comprise the total population in the problem, but they are not the categories of one of the variables. This expectation likely came from the previous questions that had populations of students and adults or middle and high school students as one of the categorical variables. Sydney said he did not know how to do this one and after I asked him what number he would put where, if he started by putting one number in a box, he randomly put 1 in the first box. He soon after requested to skip this problem. Sydney's difficulty with this problem aligns with the findings of Osterhaus et al. (2019)who suggested that problems with asymmetric variables are more difficult. Because parts (b) and (c) were based on the same problem and Sydney appeared to be quite bothered by the situation in the problem, I decided to skip these parts at this time but circled back to part b after Sydney had completed task 3. He said he would try it, questioned what it was asking, and mentioned, "This one is just not clicking in my brain." He completed part (b), which was asking for an equally likely situation, by splitting each of the given marginal frequencies in half.

More, Less, and Equally Likely

The questions in tasks 10 and 11 were asked in a different way than the previous tasks and did not use the words more, less, or equally likely, but rather asked about (in)dependence using the context of the problem. For example, tasks 10 and 11 asked:

> At Eastside High School, are current psychology students who are left-handed likely to be taller or shorter than right-handed students?
>
> Based on this information, is a box of cereal in the Evansville area more likely to be on the upper or lower shelf because it is a kids' cereal as opposed to an adults' cereal?

Students had difficulty with this wording, especially when there was a situation of independence, where a correct answer to the above questions would be no, they are neither one nor the other, but rather equally likely.

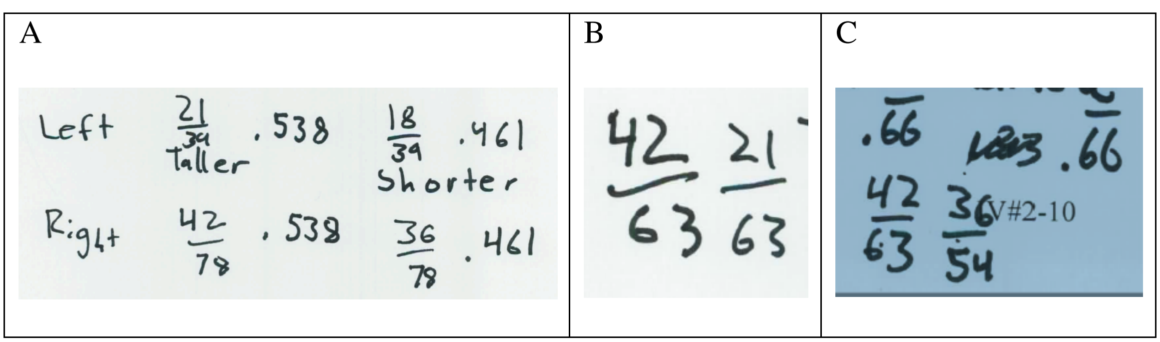
Conditioning on Columns

Tasks 9b, 10b, and 11b asked the students to condition on columns (comparing CCRFs) rather than rows (comparing RCRFs). Students adjusted to this change in different ways, with Hayden seeming to have the fewest problems. Not all participants had time to fully consider these problems; Jessie did not engage with this task at all, and Jamie had only a few minutes to work on it.

This change in direction that required comparing columns as opposed to rows was not planned in the initial protocol. Because Jamie completed tasks in less time than anticipated, I introduced part b of the problem verbally by asking a probing question to Jessie on tasks 10b and 11b. The question asked about comparing column categories (comparing CCRFs), and because Jamie had a difficult time comprehending the question, I wrote it down for him in 10b. He still struggled to understand the question and he solved it in two different ways (see Figure 38).

Figure 38

Jamie's Work Considering the Column Variable (IV#3, 10(b))



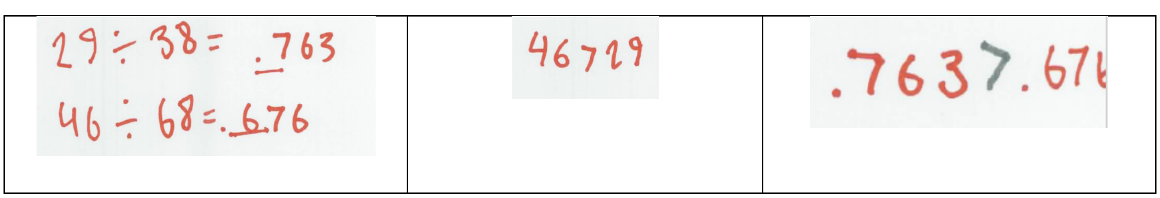
Notes. Panel A: Left and right RCRFs. Panel B: Taller CCRF. Panel C: Taller and shorter CCRFs.

First, he considered only the condition of the taller students and claimed they tended to be right-handed (L2CP). I reminded him of the question asked for a comparison with shorter students, and then he considered the taller and shorter conditions for right-handed students and noted their equality. This was about 30 minutes into the interview, so fatigue could be a factor, but I think the wording and change in the structure of the question were the reasons for Jamie's challenges. Similarly, for problem 11, Jamie computed all 4 RCRFs for part a. Part (b) asked for CCRFs, and I only asked this verbally and did not provide it in writing. Jamie responded with the same answer he gave to part a, using RCRFs. I incorporated questions aimed at comparing CCRFs in a written form into the interview protocol (9b, 10b, 11b) for all other participants.

When I asked Zander to consider column conditional relative frequencies (CCRF) for the first time (9b), he initially computed the RCRFs and then simply considered one of the columns–the patients who were cured (see Figure 39). After asking for clarification of what the question was asking and re-reading the question, then Zander reverted to considering conditional relative frequencies but did not adapt these for the columns instead of the rows. This exemplifies the difficulty in changing directions.

Figure 39

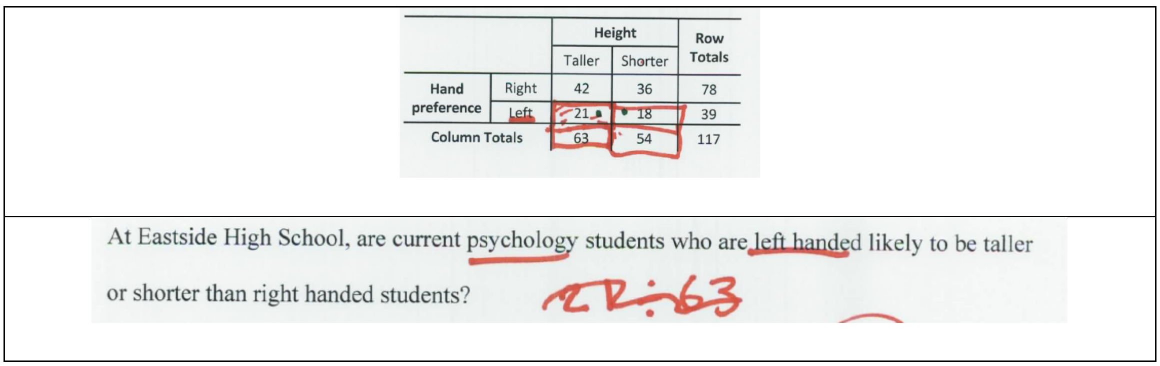
Zander's Work Considering the Column Variable (IV#3, Task 9b)



Zander struggled with the wording of problem 10, only considering the left-handed students to begin with, noting 21 of the 63 taller students were left-handed (see Figure 40). He then paused to reconsider and after re-reading the question, he mentioned there were 63 taller students and 54 shorter students, suggested 21 divided by 63 again, but then questioned it again, wondering what he should divide it by and if he should compute row or column conditional relative frequencies or the odds of a left-handed student being taller and ruling out computing the odds of a taller student being left-handed.

Figure 40

Zander's Work to Understand Question Wording (IV#3, Task 10(a))



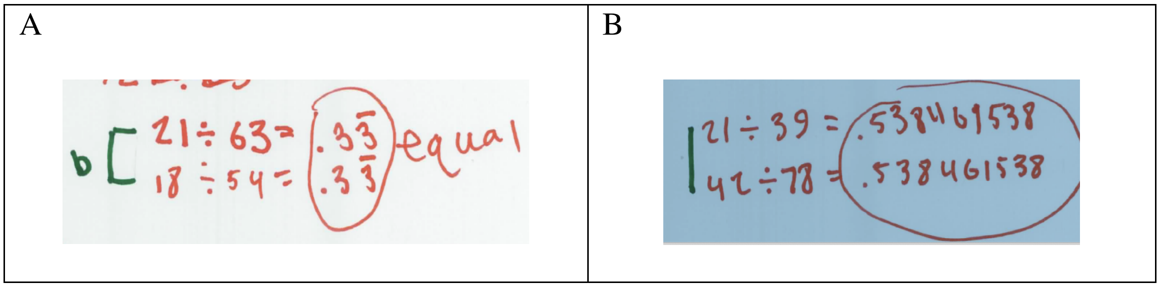
		Height		Row Totals
		Taller	Shorter	
Hand preference	Right	42	36	78
	Left	21	18	39
Column Totals		63	54	117

At Eastside High School, are current psychology students who are left handed likely to be taller or shorter than right handed students?

He proceeded with this strategy, justifying it because he was using left-handed students and comparing the difference between taller and shorter students. After computing the CCRFs using division, he concluded they were equal and then verified by computing the RCRFs (see Figure 41). Then he noticed there was not an equally likely option in the wording of the problem but seemed okay with the suggestion that an answer might be no, they are not likely to be taller or shorter. When asked which set of equations better justified the answer to the question, he said he had no idea but that since the situation was equally likely, the equations worked together.

Figure 41

Zander's Work with the Height and Handedness Task (IV#3, Task 10(a))

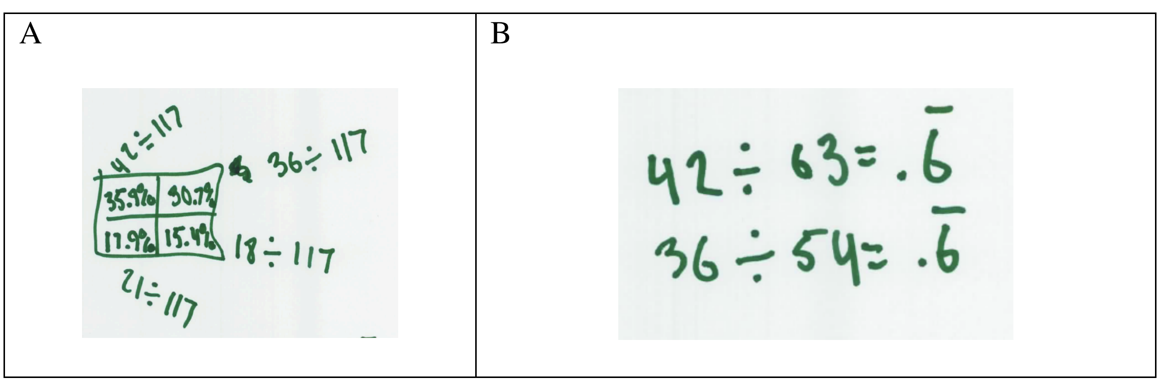


Notes. Panel A: Conditioning on columns. Panel B: Conditioning on rows.

When Zander worked with part b of this problem, which asked for column conditional relative frequencies, he initially focused on just the 63 taller students and noted there were more right-handed taller students than left-handed taller students. After being directed to re-read the question and noticing the comparison of taller and shorter students, Zander referenced the second set of equations in the first part of the problem that conditioned on the rows. He then computed the relative frequencies for all the interior cells (see Figure 42A)

Figure 42

Zander's Work with Relative Frequencies (IV#3, Task 10(b))



Notes. Panel A: Relative frequencies. Panel B: Conditional relative frequencies for right-handed students.

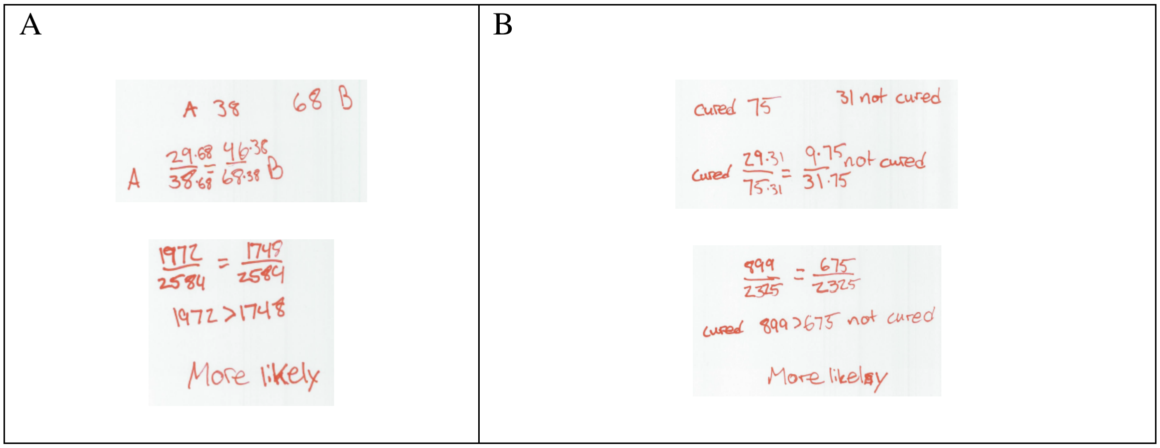
After re-reading part of the question, Zander computed the column conditional relative frequencies for the right-handed students as equal values (see Figure 42B) and requested an explanation of the question. After re-reading the question Zander was able to identify that the first set of equations from part (a) (Figure 41A) better answered this question and he noted this with a green “b” beside the equations.

When Sydney was first asked to consider the column conditional amounts in problem 9b, he read the problem several times, expressed confusion, had a moment of clarity and identified an appropriate first proportion, and then asked to skip the problem. With the next problem Sydney computed all four RCRFs in the first part, and in part b, he readily recognized he should be working with column conditional relative frequencies. When asked if the first way being equally likely implied that the second way will be equal, he said yes and justified it by saying the parts were equal and added up to the whole in the same way.

Hayden was able to adapt and apply his same approach using equivalent fractions in a different direction to the last three tasks. For tasks 9(b) and 10(b), he organized his work in a consistent way where he kept his numerical work in the middle and labeled the categories on the left and right. He used the same process where he first identified the categories of the variables he was comparing and their marginal frequencies, which become the denominator in the fractions (see Figure 43).

Figure 43

Hayden's Work Considering Row and Column Variables (IV#3, Task 9)



Notes. Panel A: Hayden's work for 9(a). Panel B: Hayden's work for 9(b).

Difficulty transferring between the two directions appeared for Hayden as he started task 11 and initially conditioned on the columns, but then realized he should condition on the rows. Hayden was older than the other students, had completed more mathematics courses, and seemed to have fewer problems with the wording of the questions. which may be the reason he had fewer problems.

Working with Mosaic Plots

Similar to students' reasoning with contingency tables, I think the way the mosaic plot questions were asked likely influenced on students' solutions. Because the questions asked about the RCRFs, it may have been more difficult for students to consider how marginal relative frequencies could be used. If different questions are asked, the results may differ. For example, using the age group and pet preference context, a question might be:

> Considering the information in the contingency table, are adults more, less, or equally likely that adults and children combined to prefer dogs as opposed to cats.

This leads students to consider the proportion of people who prefer dogs for the total number of observations.

Discussion of Findings

Participants were able to use proportional reasoning when working with contingency tables, creating mosaic plots after limited instruction, and using mosaic plots to determine (in)dependence. Students generally recognized relationships and used different approaches to calculate more efficiently. Benchmark fractions were used in splitting the total into marginal frequencies, splitting the marginal frequencies into joint frequencies, and comparing conditional relative frequencies for two categories of one variable and one category of the other variable. The use of benchmark fractions is something that might be taken into consideration as learning progresses or for differentiation. Easier tasks with complete contingency tables would include those where proportions are on either side of ½, and more difficult tasks would include proportions on the same side of ½. Tasks with complete contingency tables might require proportions to be on the same side of ½ for more difficulty. Ultimately, it is important to include situations where the proportions are on the same side of one half so that students recognize that using a benchmark fraction can identify association but does not rule out independence. Using a benchmark fraction creates an easier computational approach, but more extensive computations are required if it does not lead to a conclusion of association.

Participants in this study recognized that problems with equal marginal frequencies could be solved by comparing the whole number frequencies. It is unclear, however, how students who do not have proportional reasoning might work with these problems. Do students recognize this more efficient strategy because they can reason proportionally, or might students who do not reason proportionally be able to solve this type of problem also? Whether teachers could use tasks with contingency tables and mosaic plots to assess and further develop proportional for students requires further investigation.

Risk and odds are similar but different concepts that students often see as the same and have a difficult time understanding their differences (Ranganathan et al., 2015). These are fundamental principles to understanding measures of association such as an odds ratio and a risk ratio. Reasoning with contingency tables provides a structure where the difference between odds and risk can be explored. Participants in this study reasoned using both an odds (P3) and a risk (P1) approach. For a risk approach, they may have frequently compared CCRFs to one another rather than to an MRF because of the wording of the questions. Ideally, a student should recognize that each of these approaches (P1, P2, or P3) is valid. It is not clear from this study what younger students' innate preference is, but this information could be helpful in designing tasks for instruction. Introducing the language of odds and risk at an earlier age might help to better prepare students for further work in statistics.

There is a difference between mathematical and statistical thinking, and these two types of thinking must "work in concert"(Bargagliotti et al., 2020, p. 6). Statistics relies on precise mathematical computations and it is concerned with randomness and variability. When considering an alternative hypothesis, we look for evidence to reject it but do not require the alternative hypothesis to be true 100% of the time. Task 5 included RCRFs that were approximately equal but not exactly equal; mathematically, the proportions were different, but statistically, the proportions were close enough to be considered equally likely. This task revealed an opportunity to differentiate between mathematical and statistical thinking. This task given as a multiple-choice question on an exam would be a poor question, possibly considered to be a trick question. On the other hand, this task would be good to solve and discuss different solutions in a classroom where the goal is to recognize the difference between mathematical and statistical thinking and develop an appreciation for variability. This raises the question of how close is close enough to still be considered equally likely, and statistics educators may need to develop some guidelines for these types of problems.

Mosaic plots allow for efficient reasoning about (in)dependence and afford connections with geometry and measurement. These representations appear to be accessible to middle and high school students and have the potential to assist with their understanding of geometry and measurement. Allowing students to create a mosaic plot before reasoning with them is an important step, and continued drawing can support advancements in understanding. Students' attitudes about drawing may influence their engagement, and tasks should be structured in ways that require students to use the mosaic plot.

CHAPTER 5

SUMMARY AND CONCLUSIONS

Summary

Every high school graduate should be a statistically literate member of society, which includes the ability to reason about the statistical association of categorical variables. With more statistics and probability included in the Pre-K–12 curriculum, it is important to understand how school age students reason about this content. The limited research that has been done with student reasoning across different representations includes only completed contingency tables. In my study I considered complete and incomplete contingency tables and addressed how students reason with contingency tables and mosaic plots. The following research questions served as a guide for my study:

1. In what ways do students reason about (in)dependence of categorical variables when using contingency tables?
2. In what ways do students use mosaic plots to reason about (in)dependence of categorical variables when using contingency tables?

To investigate how younger students reason about statistical (in)dependence of categorical variables with contingency tables and mosaic plots, I conducted a qualitative study including eight initial participants from suburban public middle and high schools. I planned a series of five clinical think-aloud interviews where the first interview assessed whether the participant had the necessary reasoning to determine (in)dependence for categorical variables. This includes proportional reasoning, basic probabilistic reasoning, and understanding the structure of a contingency table. One participant did not demonstrate the necessary skills in the first interview, and two participants dropped out after the first and second interviews; thus five participants were included in this multiple-case study.

The remaining four interviews were designed to uncover student reasoning with contingency tables and mosaic plots in light of the literature and my research questions. Interviews were videotaped, student work was retained, field notes were kept, audio files were transcribed, transcripts were augmented, and lesson graphs were created. I analyzed the data in accordance with a framework developed from a review of the literature. This analysis process was not just inductive or deductive, but rather abductive. As I applied the framework to the data, I used the data to modify the framework. The resulting framework is one that emerged from and accounts for student reasoning with mosaic plots and different types of contingency tables, complete and incomplete. The five participants (Jamie, Zander, Sydney, Hayden, and Jessie) varied in how they reasoned with contingency tables alone (IV#2, IV#4) and when accompanied by a mosaic plot (IV#3, IV#5).

Jamie used his proportional reasoning abilities when working with contingency tables, including benchmark fractions, decimals, and percentages. Like some of the other participants, he started the tasks in IV#2 with an odds reasoning approach and then primarily adopted a strategy of comparing RCRFs using fractions and decimals. Rather than noticing the marginal frequencies were equal and comparing whole numbers, Jamie compared fractions but noted he could just compare the numerators when the denominators were equal. He was willing to make estimates and conjectures, like when a benchmark fraction of 1/2 was pertinent, and then evaluate these conjectures to advance his reasoning. Jamie struggled with some wording; for example, he used the term "frequency" interchangeably with "percentage," and he had difficulty when I asked him to condition on the other variable. When working with incomplete contingency tables, Jamie used his preconceived expectations to aid his solutions when they aligned and discounted them when they were not useful. He generally started with the easiest set of numbers and was able to identify ranges of values when asked to do so. Jamie readily used the mosaic plot to determine (in)dependence for complete contingency tables. When working with incomplete contingency tables with mosaic plots, he was able to estimate proportions from the mosaic plot, and he solved the cereal task that specified all marginal frequencies (IV#4, task 4) with a trial and error approach. He was able to equate the joint frequencies with the area of each tile in the mosaic plot; however, he insisted the marginal frequency was only the horizontal length as opposed to also considering the marginal frequency to be the sum of the two areas representing the subsumed joint frequencies.

Zander similarly used his proportional reasoning abilities and began solving tasks using an odds approach when working with complete contingency tables, but he did not use benchmark fractions to determine (in)dependence until a later interview. He compared whole numbers when the marginal frequencies were equal, and when they were not equal, he frequently used division equations to compute decimal values. He started comparing proportions using a whole-to-part ratio, but after struggling to explain it in context, he used a part-to-whole ratio. When working with incomplete contingency tables, Zander challenged himself to create an initial solution that was not the easiest and was frustrated when he could not solve the cereal problem where all marginal frequencies were specified. Zander readily used the mosaic plot with complete contingency tables with an odds type of reasoning, comparing the size of one joint frequency to another. He flexibly and appropriately reasoned with lengths and areas. He attributed his ability to solve the cereal problem with the incomplete contingency table to the mosaic plot. Initially, he found the joint relative frequency as a product of the marginal relative frequencies, and with prompting he recognized the marginal relative frequency could be used as a multiplier and applied to the marginal frequency.

Sydney used proportional reasoning when working with complete contingency tables and maintained his initial odds type of reasoning for an additional task (IV#2, task 4) in comparison with Jamie and Zander. After that point, he consistently used decimals to compare RCRFs and noticed an alternative strategy of comparing joint frequencies when marginal frequencies were the same. He began to use benchmark fractions with incomplete contingency tables and was able to use the marginal relative frequency as an operator to solve the cereal task without a mosaic plot. Sydney struggled with the wording of some of the problems, asking to skip some problems when he got frustrated. Sydney was willing to work with the mosaic plot when requested and he recognized it could be used to determine (in)dependence when the marginal frequencies were equal, but he preferred using the contingency table and numbers. When working with the cereal problem for a second time with the mosaic plot, Sydney was not able to solve the problem without assistance. After I directed him to consider the mosaic plot, he recognized he could use the same approach he used earlier.

Hayden used proportional reasoning and consistently compared RCRFs using fractions to determine equivalence when working with complete and incomplete contingency tables. When comparing fractions he used benchmark fractions when possible and otherwise found common denominators. Hayden recognized ranges of numbers that could work for incomplete contingency tables and tended to precision, requiring exact equivalence for an equally likely scenario. He admittedly would not use a drawing on a test and preferred to use his familiar numerical approach as opposed to the mosaic plot. When reasoning with the mosaic plot he identified the horizontal length of the first column rectangles as the important metric to compare. Like Sydney, he thought equal marginal frequencies might be a criterion for using the mosaic plot to determine (in)dependence. Hayden struggled with the cereal task and when the mosaic plot was included, he did not initially see an alternative approach to guess and check. After I directed him to consider the proportion of cereal that was on the upper shelf, he simplified the marginal relative frequency in fraction form but did not recognize he could use the fraction as a multiplier. After I had Hayden work with the mosaic plot and I suggested percentages and multiplication, Hayden was able to compute a solution. Similar to Jamie, Hayden saw the marginal frequency only as the width of the boxes in the mosaic plot, not as the sum of the areas of the two joint frequencies.

Jessie, similar to Sydney, used an odds type of reasoning through the first four tasks with complete contingency tables. Beginning with task 5, he used a scaling-up strategy where he would determine a multiplier based on the marginal frequencies. He also recognized that comparing CCRFs was another strategy that worked and when the marginal frequencies were equal, the whole number joint frequencies could be compared.

Jessie quickly recognized the efficiency of the mosaic plot for determining (in)dependence and after using it along with the contingency table for the first couple of problems, he used the mosaic plot alone for the remaining problems. Jessie started to solve the cereal problem with the mosaic plot, but the time for the interview had exceeded an hour and he eventually asked to skip the problem.

Conclusions

The five participants were able to use proportional reasoning to determine (in)dependence with categorical data in contingency tables when questions were asked in context, included the categories of the variables, and used the words less likely, equally likely, and more likely. When reasoning with complete contingency tables, students used the features of the problem and a variety of solution strategies to reason efficiently. Students recognized that equivalent marginal frequencies are a feature of contingency tables that can allow for simpler calculations. Additionally, they used benchmark fractions to evaluate equivalence and conclude (in)dependence.

Participants were able to create a solution with incomplete contingency tables that gave a condition of (in)dependence. In addition to finding a single solution, they were generally able to identify a range of solutions that were possible. The equivalence of marginal frequencies and benchmark fractions continued to play a prominent role in student reasoning. An incomplete contingency table that specified a situation of independence and included all marginal frequencies was especially challenging for students. This might be related to an incomplete understanding of the distributive property, which is about counting things in different ways where the same multiplier can be applied to each part individually or the whole collectively. When there is independence, the marginal relative frequency is equal to the corresponding conditional relative frequencies. The questions with previous tasks led students to reason with RCRFs and not MRFs and this may have caused the students to focus on the interior values.

After I instructed participants about how to create a mosaic plot, they were able to create one for given contingency tables without assistance. Additionally, they were able to use mosaic plots alone and alongside contingency tables to reason about (in)dependence. Students readily recognized that a vertical line that cuts through both rows signifies independence and that this line does not have to be in the middle for the situation to be equally likely. The relevance of the horizontal line being in the middle was more challenging, and some students thought this was a necessary condition for using the mosaic plot to reason about (in)dependence. This might be due to the prominence of benchmark fractions and an incomplete understanding of their application. Students' attitudes about drawings may have contributed to this. Benchmark fractions continued to be helpful and students tended to the areas and linear distances when reasoning with mosaic plots.

Reasoning with mosaic plots and incomplete contingency tables afforded students multiple representations for their reasoning. Students who were more open to using drawings for mathematical reasoning made better connections between a mosaic plot and a contingency table. Creating a mosaic plot required attention to the measured quantities. Continued drawing of mosaic plots allowed students to reason more deeply about the relationships among the measured quantities because they had material to conceive quantities, their measures, and their relationships. Students were able to use a mosaic plot to estimate the frequencies in a contingency table when given unequal row marginal frequencies. This required estimating proportions from each row of the mosaic plot. For a challenging problem with only the marginal frequencies and a condition of independence, more students found a solution when a mosaic plot accompanied the contingency table. A mosaic plot allows students to visually recognize the distributive property. Students suggested that including a grid and numbers on the mosaic plot would make it more useful.

Participants were challenged when asked to condition on the column variable rather than the row variable. This was after they had answered several questions asking to condition on the row variable, including a question with the same content. This indicates that changing the direction of conditioning is challenging but does not compare one direction to the other. Some of the wording of the problems presented difficulties for the students in determining what proportions were important to consider. Students had understandings of the words independent and association that might have created difficulties in understanding (in)dependence in the context of categorical data in the problems, and their understanding of a variable was limited.

Implications

In the brief amount of time I worked with the participants, I saw how some aspects of contingency tables impacted student reasoning. It is important to consider how students reason when educating prospective teachers, conducting professional development, or developing curriculum. The findings from my study inform these activities in a variety of ways.

Question Wording

The wording of questions can be challenging and understanding what the questions are asking can be difficult. Question wording has implications for developing instructional sequences. The questions I asked specified categories of the variables whereas Batanero used questions that asked about the variables in general and allowed any combination of relative frequencies to be compared. The questions I asked might be placed earlier in instruction that is designed based on a simple-to-complex sequence. Instructional design that includes consideration of pre-requisite skills and student choice might use these criteria to determine which type of question is asked. The questions I asked that were based on the categories might also be used as a formative assessment or to evaluate students' pre-requisite skills.

Additionally, the questions I asked likely led the students to reason by comparing conditional relative frequencies (P1) as opposed to comparing a conditional relative frequency to a marginal relative frequency (P2). Question wording may be a significant factor in how students reason, and it is reasonable to think it is an important factor to consider when developing an instructional sequence. It is important to include different types of questions, especially those that consider conditional marginal frequencies, because this comparison becomes more important as the dimensions of tables expand.

The questions I asked generally specified all four categories for the two variables, which may have made it more difficult to select which variable to condition on. For example, one of the cereal tasks (IV#2 and IV#3, task 11a) asked students to consider both the upper and lower shelf and the kids' and adults' cereal and condition on the type of cereal (see Figure 44). Alternatively, if the question asked the students to consider only the cereal on the upper shelf and compare the kids' cereal to the adults' cereal, it might be easier to conclude that the different categories of the cereal type are the conditions that should be compared.

Figure 44

Conditioning on the Row Variable

Condition

		Cereal location		Row Totals
		Upper shelf	Lower shelf	
Cereal type	Kids' cereal	62	47	109
	Adults' cereal	33	17	50
Column Totals		95	64	159

Compare

62/109 = 57%

33/50 = 66%

95/159 = 60%

Similarly, when conditioning on the column variable, the cereal location in this example, a question might specify only one of the categories for the cereal type, the kids' cereal (see Figure 45). This might be helpful when a series of tasks are asking to condition on different variables.

Figure 45

Conditioning on the Column Variable

Condition

		Cereal location		Row Totals
		Upper shelf	Lower shelf	
Cereal type	Kids' cereal	62	47	109
	Adults' cereal	33	17	50
Column Totals		95	64	159

Compare

62/95 = 65% 47/64 = 73% 95/159 = 69%

The wording of questions can also be addressed by having students create questions, which allows teachers to formatively assess students. This may provide students an opportunity to develop mathematically and statistically appropriate language and better understand contingency tables.

Design aspects

It is important to consider different aspects of contingency tables when designing instruction. For example, problems with a clear explanatory and response variable, population comparisons, and symmetric variables might be easier for students. Instructional sequences should consider these aspects as well, and the numbers should be selected based on real or realistic context and in consideration of the ways students reason. Because students rely on benchmark fractions, problems should include those representing dependence that require comparison of proportions that are on the same side of ½. For a simple-to-complex instructional sequence, this might begin with problems with the same marginal frequency so they can compare whole numbers and extend to problems with different marginal frequencies so they need to compare proportions.

Ultimately, it is important that students reason flexibly comparing either row or column conditional relative frequencies to each other or the corresponding marginal relative frequency. Because there are many constituent components to coordinate, instruction might begin with a standard structure where the explanatory variable is in the same place, either the row or column variable. Initial questions can lead students to condition on a row or a column variable. Once a solid solution strategy is adopted, problems can be extended to consider the other variable and more difficult aspects of contingency tables.

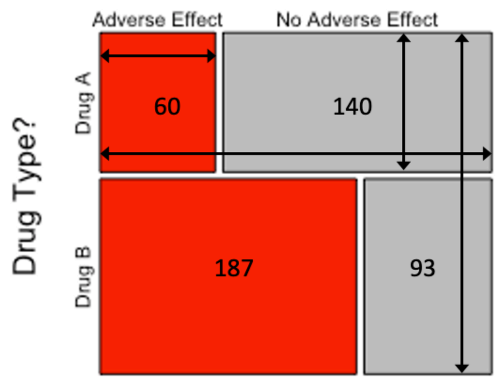
Mosaic Plots

Mosaic plots are useful when reasoning about (in)dependence for categorical variables. Students were able to create and reason with sideways mosaic plots that were in a horizontal orientation aligning with the contingency table and the question posed. Reasoning with multiple representations requires students to make more connections and helps students to reason more deeply about concepts (Dreher & Kuntze, 2015). Because the students created the mosaic plots, they made explicit connections between them and the contingency table. Mosaic plots used in conjunction with contingency tables allowed some students to solve problems they could not solve with the contingency table alone.

Mosaic plots have a clear connection with geometry, and students tended to different geometric features when reasoning with mosaic plots. Because geometry is concurrently studied in middle and high school, tasks connecting a mosaic plot and geometry concepts are appropriate. The frequencies in a contingency table have different connections to one-dimensional and two-dimensional measurements of the different tiles in a mosaic plot. For example, consider a mosaic plot that is labeled with joint frequencies and delineates the horizontal and vertical distances that might be considered when comparing conditional relative frequencies (see Figure 46). Many different questions can be posed, and different geometric terms can be used. This provides an opportunity for students to tend to precision and distinctly note the difference between horizontal and vertical as well as one-dimension and two-dimension measurements.

Figure 46

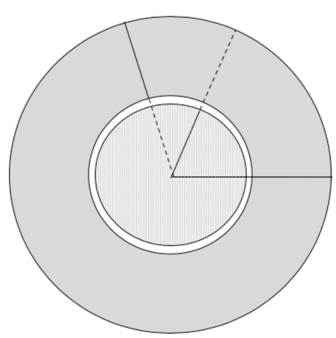
Mosaic Plot with Labels



A more advanced question might ask students to consider areas. For example, what are the dimensions of each rectangle? Additionally, the students could measure the heights and lengths in centimeters or inches and notice how the same relationships hold. Furthermore, the differences in student answers could be used to address measurement variability. Additional geometry connections could be made by extending the design principles of a mosaic plot to a circle where the areas of the inner circle and outer ring are proportional to the marginal distribution (see Figure 47). Then either the areas or the angle measures can be used to compare proportions.

Figure 47

Circular Mosaic Plot



Other Representations

Other representations can be used when reasoning with categorical data. These include Venn diagrams and various sorts of bar graphs (e.g., side-by-side and segmented bar graphs with scales of frequencies and relative frequencies). Students were able to create bar graphs without assistance but, they had difficulty using Venn diagrams without instruction. Because participants struggled to clearly identify a variable and category for a contingency table, a Venn diagram, which requires one circle to be identified as a category of one variable and the other circle to be identified as a category of the other variable, might prove to be useful. Different representations can help with different areas of understanding and we need to allow students to create and reason across multiple representations.

Statistical Reasoning

Tending to precision mathematically seemed to override an appreciation for variability. Statistical reasoning is different than mathematical reasoning, and variability is a concept that takes time to develop. We need to introduce statistical reasoning earlier in the curriculum so that students develop an appreciation for when precision is important and when variability is important. These students were able to engage in sophisticated reasoning with categorical variables, so we might consider enhancing rather than reducing attention to statistics in the curriculum. Reasoning about the association of categorical variables using a contingency table provides connections to real-world contexts such as medicine, agriculture, retail, surveys, etc. and students can easily create real data through observations.

Limitations

I faced limitations in my study related to the participants and the design. Qualitative research, whereas not generalizable, can be generative and transferrable (Lincoln & Guba, 1986). These results offer insight into how students reason about contingency tables and mosaic plots and can be considered groundwork for future research. Another limitation related to the participants is that they were all boys from suburban middle and high schools in a higher track of mathematics classes at schools that perform well above-average. Ideally, I would have liked the participants to be more varied in many of these characteristics. I did not create criteria about gender or math placement for participation and rather placed more importance on finding participants who I knew through my community and I thought might perform well in a think-aloud interview. I requested participation from and recruited more girls, but they were either too busy, dropped out, or did not pass the screening criteria in the first interview.

When developing the protocols and determining probing questions, I did not know the mathematical background and performance of my participants. I anticipated alternative solution methods and used the previous interviews with participants to inform the protocols. The participants were interacting with me while I was in the position of a researcher, and this could have influenced how they responded. The setting included video cameras and was not in a place where they routinely worked on math problems, which may have influenced their work. Participants' past experiences in math classes focused on getting the correct answer, and this may have interfered with authentic think-aloud reasoning. To account for some of these limitations, I tried to make the participants as comfortable as possible, reminding them there were no wrong answers and I was not interested in a correct answer, but rather how they were thinking. I tried to ask students what they were thinking, provide wait time for additional thoughts to emerge, and question myself as to whether they revealed their thinking before asking more probing questions.

When designing tasks for the four interviews that addressed reasoning with contingency tables and mosaic plots, I did not include tasks that addressed all components of the statistical problem-solving process. Additionally, the tasks required the participants to compute the proportions rather than providing them. Thus, the results focus more on students' application of mathematics and reasoning across representations rather than their understanding of statistics. I was not able to address how all the different aspects of a contingency table might impact reasoning. Ideally, variations of all aspects would be addressed and one aspect at a time might be varied to gain an understanding of how students respond. Because I had a limited number of tasks I could complete in four interviews, I selected tasks with aspects that were less likely to make reasoning more difficult.

Future Research

This study extended past work in this area by having students reason with incomplete contingency tables and mosaic plots after having instruction on how to create them. This created a deeper understanding of student reasoning and uncovered connections to geometry; yet there is still much work to be done in this area. It is important to understand this content, not only in an isolated problem-solving task but within the scope of the statistical problem-solving process. Thus, future studies should be designed with this in mind.

There are many aspects of contingency tables addressed in Chapter 2, and future work might systematically consider their impact on student reasoning. Some of the aspects I identified have not been addressed (e.g., population comparison vs. association of variables, explanatory and response variables) whereas others have limited studies (e.g., symmetric and asymmetric, table dimension & size). More insights into how students reason through the statistical problem-solving process as these different aspects are considered can inform learning trajectories as well as identify connections with other mathematical content.

We live in a multi-variate world and variables often have more than two categories. Limited studies have considered contingency tables beyond the simplest 2x2 table. Future studies might consider two-way tables with variables that have more than two categories (e.g., 2x3, 3x2, 3x3). This transition parallels the progression to considering three variables in algebra or three dimensions in geometry. The transition from two-dimensions to three-dimensions is more of a challenge for students than the transition from one to two variables (Yerushalmy, 1997). Future studies might consider contingency tables for more than two variables (e.g., 2x2x2, 2x2x3, 2x3x4).

Reasoning with representations that do not include numbers invokes quantitative reasoning. Mosaic plots contain additional information that is not included in a contingency table with just frequencies because the proportions are presented visually. Future research might first provide only a mosaic plot for a situation, posing questions about the relationships of the variables. Alternatively, including the percentages with a contingency table, similar to statistical software output, puts the two representations on the same footing. Including tasks with this additional information could change the focus from mathematical calculations and elucidate the relationships.

A variety of representations can be used (e.g., segmented bar charts, side-by-side bar charts, mosaic plots, sieve plots, etc.), and these can be labeled in different ways (e.g., frequencies, relative frequencies, decimals, percentages, etc.). Graphical representations can be displayed in different directions and depending on the variable that is being conditioned on, they may or may not align with the contingency table. It is reasonable to expect this might be an important consideration, especially for students who have not developed spatial reasoning. Future research should consider different representations, their strengths in aiding understanding, and how students reason across them.

In Closing

Through working with middle and high school students, I recognize their desire to understand the underlying mathematics for a given task and to reach conclusions efficiently. They are in the process of developing their vocabulary, and mathematical terms not used in everyday language, or used in a different manner, are challenging. Their explanations may sometimes seem incorrect or incomplete because of imprecise vocabulary or connections, but they seem to have logical underpinnings based on connections they have made or are forming. I found that students were very concerned about getting the right answer, and the fact these interviews were not concerned with correct answers seemed simultaneously refreshing and taxing to the participants: refreshing in the sense that someone was interested in what they thought, not just how they performed against a standard; taxing in a sense that thinking-aloud while completing a task is not what they were used to and requires a lot of attention. I found that students applied proportional reasoning with contingency tables and even though some students found drawing representations to be more juvenile, they were all able to create a mosaic plot and use it to recognize (in)dependence, or as they thought of it in most of these problems, when it is “equally likely.” Reasoning with multiple representations deepens students conceptual understanding and reasoning with categorical variables is an important skill for members of society to have.

REFERENCES

Agresti, A. (2007). *An introduction to categorical data analysis* (2nd ed). Wiley-Interscience. https://doi.org/10.1002/0470114754

Agresti, A., & Franklin, C. (2009). *Statistics: The art and science of learning from data* (Second). Pearson Prentice Hall.

Arnold, P. (2008). *What about the P in the PPDAC cycle? An initial look at posing questions for statistical investigation* [Paper Presentation]. Eleventh International Congress of Mathematics Education (ICME-11), Monterrey, Mexico.

Bargagliotti, A., Franklin, C. A., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D. A. (2020). *Pre-K—12 guidelines for assessment and instruction in statistics education II (GAISE II)* (Second). American Statistical Association.

Batanero, C., Cañadas, G. R., Diaz, C., & Gea, M. M. (2015). Judgment of association between potential factors and associated risk in 2x2 tables: A study with psychology students. *The Mathematics Enthusiast*, *12*(1), 347—363.

Batanero, C., Cañadas, G. R., Estepa, A., & Arteaga, P. (2012). Psychology students' estimation of association. In L. R. Van Zoest, J.-J. Lo, & J. L. Kratky (Eds.), *Proceedings of 36th Conference of the International Group for Psychology in Mathematics Education* (Vol. 2, p. 51—58). Western Michigan University.

Batanero, C., Estepa, A., Godino, J. D., & Green, D. R. (1996). Intuitive strategies and preconceptions about association in contingency tables. *Journal for Research in Mathematics Education*, *27*(2), 151—169. https://doi.org/10.2307/749598

Batanero, C., & Godino, J. D. (1998). Understanding graphical and numerical representations of statistical association in a computer environment. In L. Pereira-Mendoza (Ed.), *Proceedings of the Fifth International Conference on Teaching Statistics* (Vol. 2, p. 1017—1024). International Statistical Institute. https://iase-web.org/Conference_Proceedings.php?p=ICOTS_5_1998

Batanero, C., Godino, J. D., & Estepa, A. (1998). Building the meaning of statistical association through data analysis activities. In A. Olivier & K. Newstead (Eds.), *PME 22: Proceedings of the 22nd conference of the International Group for the Psychology of Mathematics Education* (p. 248—269). Citeseer.

Batanero, C., & Sanchez, E. (2005). What is the nature of high school students' conceptions and misconceptions about probability? In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (Vol. 40, p. 241—266). Springer. https://doi.org/10.1007/0-387-24530-8_11

Beckmann, S. (2018). *Mathematics for elementary teachers, with activities* (5th edition). Pearson.

Beckmann, S., & Izsák, A. (2015). Two perspectives on proportional relationships: Extending complementary origins of multiplication in terms of quantities. *Journal for Research in Mathematics Education*, *46*(1), 17—38. https://doi.org/10.5951/jresematheduc.46.1.0017

Behr, M. J., Harel, G., Post, T. R., & Lesh, R. (1992). *Rational number, ratio, and proportion*. (D. A. Grouws, Ed.). Macmillan Publishing.

Bell, A., & Burkhardt, H. (2002). *Domain frameworks in mathematics and problem solving* [Paper Presentation]. Annual meeting of the American Educational

Research Association, New Orleans, LA. www. nottingham. ac. uk/education/MARS/papers

Benson, C. C., Wall, J. J., & Malm, C. (2013). The distributive property in grade 3? *Teaching Children Mathematics*, *19*(8), 498–506. https://doi.org/DOI: 10.5951/teacchilmath.19.8.0498

Boaler, J., Chen, L., Williams, C., & Cordero, M. (2016). Seeing as understanding: The importance of visual mathematics for our brain and learning. *Journal of Applied & Computational Mathematics*, *5*(5), 1–6. https://doi.org/10.4172/2168-9679.1000325

Casey, S. A., Albert, J., & Ross, A. (2018). Developing knowledge for teaching graphing of bivariate categorical data. *Journal of Statistics Education*, *26*(3), 197–213. https://doi.org/10.1080/10691898.2018.1540915

Casey, S. A., Hudson, R., & Ridley, L. (2018). Students' reasoning about association of categorical variables. In M. A. Sorto, A. White, & L. Guyot (Eds.), *Looking Back, Looking Forward: Proceedings of the Tenth International Conference on Teaching Statistics*. International Statistical Institute. http://iase-web.org/Conference_Proceedings.php?p=ICOTS_10_2018

Casey, S. A., Ross, A. M., Groth, R. E., & Zejullahi, R. (2015). Developing teachers' knowledge of content and students for teaching categorical association. In T. G. Bartell, K. N. Bieda, P. P. Putnam, K. Bradfield, & H. Dominguez (Eds.), *37th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 872–875). Michigan State University.

Charters, E. (2003). The use of think-aloud methods in qualitative research An introduction to think-aloud methods. *Brock Education Journal*, *12*(2), 68–82. https://doi.org/10.26522/brocked.v12i2.38

Cherry, W. H., & Oldford, R. W. (2002). *Eikosograms for teaching probabilistic independence and its modelling*. [Unpublished manuscript]. http://www.stats.uwaterloo.ca/~rwoldfor/papers/independence/paper.pdf

Clark, M. R., Berenson, S. B., & Cavey, L. O. (2003). A comparison of ratios and fractions and their roles as tools in proportional reasoning. *The Journal of Mathematical Behavior*, *22*(3), 297–317. https://doi.org/10.1016/S0732-3123(03)00023-3

Clarke, D. M., & Roche, A. (2009). Students' fraction comparison strategies as a window into robust understanding and possible pointers for instruction. *Educational Studies in Mathematics*, *72*(1), 127–138. https://doi.org/10.1007/s10649-009-9198-9

Clement, J. (2000). Analysis of clinical interviews: Foundations and model viability. In A. E. Kelley & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 547–589). Lawrence Erlbaum Associates.

Cobb, G. W., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *The American Mathematical Monthly*, *104*(9), 801–823. https://doi.org/10.1080/00029890.1997.11990723

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed). SAGE Publications.

Curry, M., & Outhred, L. (2005). Conceptual understanding of spatial measurement. In P. Clarkson, A. Downton, D. Gronn, A. McDonough, R. Pierce, & A. Roche (Eds.), *Building connections: Theory, research and practice*. *Proceedings of the 28th annual conference of the Mathematics Education Research Group of Australasia* (p. 265—272). MERGA.

Davidian, M., & Louis, T. A. (2012). Why statistics? *Science*, *336*(6077), 12–12. https://doi.org/10.1126/science.1218685

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2005). *The SAGE handbook of qualitative research* (3rd ed). SAGE Publications.

Ding, M., & Li, X. (2010). A comparative analysis of the distributive property in US and Chinese elementary mathematics textbooks. *Cognition and Instruction*, *28*(2), 146–180. https://doi.org/10.1080/07370001003638553

diSessa, A. A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and Instruction*, *22*(3), 293–331. https://doi.org/10.1207/s1532690xci2203_2

Dougherty, B., Bryant, D. P., Bryant, B. R., & Shin, M. (2016). Helping students with mathematics difficulties understand ratios and proportions. *TEACHING Exceptional Children*, *49*(2), 96–105. https://doi.org/10.1177/0040059916674897

Dreher, A., & Kuntze, S. (2015). Teachers' professional knowledge and noticing: The case of multiple representations in the mathematics classroom. *Educational Studies in Mathematics*, *88*(1), 89–114. https://doi.org/10.1007/s10649-014-9577-8

Fischbein, E., Deri, M., Nello, M. S., & Marino, M. S. (1985). The role of implicit models in solving verbal problems in multiplication and division. *Journal for Research in Mathematics Education*, *16*(1), 3—17. https://doi.org/10.2307/748969

Fisher, L. C. (1988). Strategies used by secondary mathematics teachers to solve proportion problems. *Journal for Research in Mathematics Education*, 157–168.

Franklin, C., Mewborn, D., & Burrill, G. F. (2006). The statistical education of Pre-K—12 teachers: A shared responsibility. In *Thinking and reasoning with data and chance: Sixty-eighth annual yearbook of the National Council of Teachers of Mathematics* (p. 335—344). National Council of Teachers of Mathematics.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: A pre-K—12 curriculum framework*. American Statistical Association.

Friel, S. N., Curcio, F. R., & Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, *32*(2), 124. https://doi.org/10.2307/749671

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, *70*(1), 1–25. https://doi.org/10.1111/j.1751-5823.2002.tb00336.x

Garcia-Retamero, R., & Hoffrage, U. (2013). Visual representation of statistical information improves diagnostic inferences in doctors and their patients. *Social Science & Medicine*, *83*, 27–33. https://doi.org/10.1016/j.socscimed.2013.01.034

Gibson, B. (1997). Talking the test: Using verbal report data in looking at the processing of cloze Tasks. *Edinburgh Working Papers In Applied Linguistics*, *8*, 54–62.

Ginsburg, H. (1981). The clinical interview in psychological research on mathematical thinking: Aims, rationales, techniques. *For the Learning of Mathematics*, *1*(3), 4–11.

Ginsburg, H. (1997). *Entering the child's mind: The clinical interview in psychological research and practice*. Cambridge University Press.

Glencross, M. (1998). Building the meaning of statistical association through data analysis activities. A reaction to Batanero et. Al. In A. Olivier & K. Newstead (Eds.), *PME 22: Proceedings of the 22nd conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, p. 264—242).

Godino, J. D., Batanero, C., & Font, V. (2007). The onto-semiotic approach to research in mathematics education. *ZDM*, *39*(1–2), 127–135. https://doi.org/10.1007/s11858-006-0004-1

Hall, R. (2000). Videorecording as theory. In A. E. Kelley & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 647–664). Lawrence Erlbaum Associates, Inc.

Harel, G., & Behr, M. (1995). Teachers' solutions for multiplicative problems. *Hiroshima Journal of Mathematics Education*, *3*, 31–51.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking: From childhood to adolescence*. (A. Parsons & S. Milgram, Trans.). Basic Books. https://doi.org/10.1037/10034-000

Izsák, A. (2003). “ We want a statement that is always true”: Criteria for good algebraic representations and the development of modeling knowledge. *Journal for Research in Mathematics Education*, *34*(3), 191—227. https://doi.org/10.2307/30034778

Izsák, A., & Jacobson, E. (2017). Preservice teachers’ reasoning about relationships that are and are not proportional: A Knowledge-in-Pieces Account. *Journal for Research in Mathematics Education*, *48*(3), 300—339. https://doi.org/10.5951/jresematheduc.48.3.0300

Jewitt, C. (2012). *An introduction to using video for research* (An Introduction to Using Video for Research, p. 26) [Institute of Education Working Paper]. National Centre for Research Methods. http://eprints.ncrm.ac.uk/2259/4/NCRM_workingpaper_0312.pdf

Kahneman, D., & Tversky, A. (1982). On the study of statistical intuitions. *Cognition*, *11*(2), 123–141. https://doi.org/10.1016/0010-0277(82)90022-1

Kaput, J. J., & West, M. M. (1994). Missing value proportional reasoning problems: Factors affecting informal reasoning patterns. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics* (p. 235—287). State University of New York Press.

Karp, K. S., Bush, S. B., & Dougherty, B. J. (2015). 12 math rules that expire in the middle grades. *Mathematics Teaching in the Middle School*, *21*(4), 208. https://doi.org/10.5951/mathteacmiddscho.21.4.0208

Karplus, R., Pulos, S., & Stage, E. K. (1983a). Early adolescents' proportional reasoning on 'rate'problems. *Educational Studies in Mathematics*, *14*(3), 219–233. https://doi.org/10.1007/BF00410539

Karplus, R., Pulos, S., & Stage, E. K. (1983b). Proportional reasoning of early adolescents. In R. Lesh & M. Landau (Eds.), *Acquisition of mathematics concepts and processes* (pp. 45–90). Academic Press.

Kuiper, S. (2010, July). Incorporating a research experience into an early undergraduate statistics course. *International Conference on Teaching Statistics (ICOTS)*.

Kuusela, H., & Pallab, P. (2000). A comparison of concurrent and retrospective verbal protocol analysis. *The American Journal of Psychology*, *113*(3), 387.

Lamon, S. J. (1993). Ratio and proportion: Connecting content and children's thinking. *Journal for Research in Mathematics Education*, *24*(1), 41–61. https://doi.org/10.2307/749385

Lamon, S. J. (2007). Rational numbers and proportional reasoning: Toward a theoretical framework for research. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 1, pp. 629–667). Information Age Publishing.

Lamon, S. J. (2012). *Teaching fractions and ratios for understanding: Essential content knowledge and instructional strategies for teachers* (3rd ed.). Routledge. https://doi.org/10.4324/9780203803165

Lehrer, R. (2012). Longitudinal study of children's reasoning about space and geometry. In R. Lehrer & D. Chazan (Eds.), *Designing learning environments for* *developing understanding of geometry and space* (pp. 137–167). Lawrence Erlbaum Associates. https://doi.org/10.4324/9780203053461-12

Lincoln, Y. S., & Guba, E. G. (1986). But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. *New Directions for Program Evaluation*, *1986*(30), 73–84. https://doi.org/10.1002/ev.1427

Lobato, J., & Ellis, A. (2010). *Developing essential understanding of ratios, proportions, and proportional reasoning for teaching mathematics: Grades 6—8*. (R. Charles, Ed.). National Council of Teachers of Mathematics.

Marshall, C., & Rossman, G. B. (2016). *Designing qualitative research* (Sixth edition). Sage.

Maxwell, J. A. (2013). *Qualitative research design: An interactive approach* (3rd ed). SAGE Publications.

Merriam, S. B. (1998). *Qualitative research and case study applications in education* (2nd ed). Jossey-Bass Publishers.

Moore, K. C. (2014). Quantitative reasoning and the sine function: The case of Zac. *Journal for Research in Mathematics Education*, *45*(1), 102–138. https://doi.org/10.5951/jresematheduc.45.1.0102

National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematical success for all*. National Council of Teachers of Mathematics.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards for mathematics*. National Governors Association Center for Best Practices, Council of Chief State School Officers, Washington D.C.

Noelting, G. (1980). The development of proportional reasoning and the ratio concept part II: Problem-structure at successive stages; Problem-solving strategies and the mechanism of adaptive restructuring. *Educational Studies in Mathematics*, *11*(3), 331–363. https://doi.org/10.1007/BF00697744

Nunes, T., Bryant, P., Evans, D., Gottardis, L., & Terlektsi, M.-E. (2014). The cognitive demands of understanding the sample space. *ZDM*, *46*(3), 437–448. https://doi.org/10.1007/s11858-014-0581-3

Obersteiner, A., Bernhard, M., & Reiss, K. (2015). Primary school children's strategies in solving contingency table problems: The role of intuition and inhibition. *ZDM: The International Journal on Mathematics Education*, *47*(5), 825–836. https://doi.org/DOI 10.1007/s11858-015-0681-8

Obersteiner, A., Reiss, K., & Bernhard, M. (2016). How do primary school children solve contingency table problems that require multiplicative reasoning? *Proceedings of the 40th Conference of the International Group for the Psychology of Mathematics Education*, *3*, 387–394.

Oldford, R. W., & Cherry, W. H. (2006). *Picturing probability: The poverty of Venn diagrams, the richness of eikosograms* [Unpublished manuscript]. http://sas.uwaterloo.ca/~rwoldfor/papers/venn/eikosograms/paperpdf.pdf

Orrill, C. H., & Brown, R. E. (2012). Making sense of double number lines in professional development: Exploring teachers' understandings of proportional relationships. *Journal of Mathematics Teacher Education*, *15*(5), 381–403.

Osterhaus, C., Magee, J., Saffran, A., & Alibali, M. W. (2019). Supporting successful interpretations of covariation data: Beneficial effects of variable symmetry and problem context. *Quarterly Journal of Experimental Psychology*, *72*(5), 994–1004. https://doi.org/10.1177/1747021818775909

Park, J., & Brannon, E. M. (2013). Training the approximate number system improves math proficiency. *Psychological Science*, *24*(10), 2013–2019. https://doi.org/10.1177/0956797613482944

Patton, M. Q. (2005). *Qualitative research* (3rd ed.). Sage.

Patton, M. Q., & Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3 ed). SAGE Publications.

Pfannkuch, M., & Budgett, S. (2017). Reasoning from an eikosogram: An exploratory study. *International Journal of Research in Undergraduate Mathematics Education*, *3*(2), 283–310. https://doi.org/10.1007/s40753-016-0043-0

Piaget, J., & Inhelder, B. (1975). *The origin of the idea of chance in children*. (L. Leake, Jr., P. Burrell, & H. D. Fishbein, Trans.). W. W. Norton & Company, Inc.

Ranganathan, P., Aggarwal, R., & Pramesh, C. S. (2015). Common pitfalls in statistical analysis: Odds versus risk. *Perspectives in Clinical Research*, *6*(4), 222–224. https://doi.org/10.4103/2229-3485.167092

Riehl, S. M., & Steinthorsdottir, O. B. (2014). Revisiting Mr. Tall and Mr. Short. *Mathematics Teaching in the Middle School*, *20*(4), 220. https://doi.org/10.5951/mathteacmiddscho.20.4.0220

Saffran, A., Barchfeld, P., Sodian, B., & Alibali, M. W. (2016). Children's and adults' interpretation of covariation data: Does symmetry of variables matter? *Developmental Psychology*, *52*(10), 1530–1544. https://doi.org/10.1037/dev0000203

Saldaña, J. (2015). *The coding manual for qualitative researchers*. SAGE Publications.

Schoenfeld, A. H. (2002). Making mathematics work for all children: Issues of standards, testing, and equity. *Educational Researcher*, *31*(1), 13–25. https://doi.org/10.3102/0013189X031001013

Seago, N. (2003). Using video as an object of inquiry for mathematics teaching and learning. In J. Brophy (Ed.), *Using Video in Teacher Education* (Vol. 10, pp. 259–286). Emerald (MCB UP). https://doi.org/10.1016/S1479-3687(03)10010-7

Seidman, I. (2013). *Interviewing as qualitative research: A guide for researchers in education and the social sciences*. Teachers College Press.

Shaklee, H., & Mims, M. (1981). Development of rule use in judgments of covariation between events. *Child Development*, *52*(1), 317. https://doi.org/10.2307/1129245

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning*. NCTM, Information Age Publishing; 957-1009.

Siegler, R. S. (2017, November 28). Fractions: Where it all goes wrong. *Scientific American*. https://www.scientificamerican.com/article/fractions-where-it-all-goes-wrong/

Stake, R. E. (1995). *The art of case study research*. SAGE Publications.

Thompson, P. W. (1994). The development of the concept of speed and its relationship to concepts of rate. In Guershon Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics* (pp. 179–234). SUNY Press.

Tourniaire, F., & Pulos, S. (1985). Proportional reasoning: A review of the literature. *Educational Studies in Mathematics*, *16*(2), 181–204. https://doi.org/10.1007/PL00020739

Usiskin, Z., & Hall, K. (2015). The relationships between statistics and other subjects in the K—12 curriculum. *Chance*, *28*(3), 4–18. https://doi.org/10.1080/09332480.2015.1099361

Van de Walle, J. A. (1994). *Elementary and middle school mathematics: Teaching developmentally*. Addison Wesley Longman.

Van de Walle, J. A., Bay-Williams, J. M., Lovin, L. H., & Karp, K. S. (2018). *Teaching student-centered mathematics. Developmentally appropriate instruction for grades 6—8* (Third edition). Pearson.

Van de Walle, J. A., Karp, K., Bay-Williams, J. M., & Wray, J. A. (2019). *Elementary and middle school mathematics: Teaching developmentally* (Tenth edition). Pearson.

Watson, J. M., & Callingham, R. (2014). Two-way tables: Issues at the heart of statistics and probability for students and teachers. *Mathematical Thinking and Learning*, *16*(4), 254–284. https://doi.org/10.1080/10986065.2014.953019

Watson, J. M., & Callingham, R. (2015). Lung disease, indigestion, and two-way tables. *Investigations in Mathematics Learning*, *8*(2), 1–16. https://doi.org/10.1080/24727466.2015.11790348

Watson, J. M., Callingham, R. A., & Kelly, B. A. (2007). Students' appreciation of expectation and variation as a foundation for statistical understanding.

Mathematical Thinking and Learning, *9*(2), 83–130. https://doi.org/10.1080/10986060709336812

Watson, J. M., & Moritz, J. B. (2002). School students' reasoning about conjunction and conditional events. *International Journal of Mathematical Education in Science and Technology*, *33*(1), 59–84.

Yazan, B. (2015). Three approaches to case study methods in education: Yin, Merriam, and Stake. *The Qualitative Report*, *20*(2), 134–152.

Yeo, J. B. W. (2017). Development of a framework to characterise the openness of mathematical tasks. *International Journal of Science and Mathematics Education*, *15*(1), 175–191. https://doi.org/10.1007/s10763-015-9675-9

Yerushalmy, M. (1997). Designing representations: Reasoning about functions of two variables. *Journal for Research in Mathematics Education*, *28*(4), 431–466. JSTOR. https://doi.org/10.2307/749682

Yin, R. K. (2018). *Case study research and applications: Design and methods* (Sixth edition). SAGE Publications.

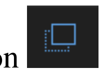
APPENDICES

APPENDIX A

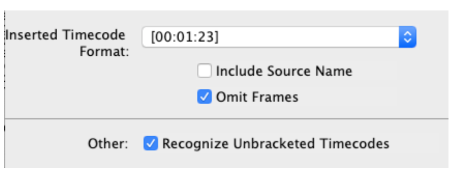
INSTRUCTIONS FOR INTERVIEW DATA MANAGEMENT

1. Scan in written work
    - Scan to USB on printer (scan-up arrow-change settings- to PDF – color)
    - Copy to folder with the name IV#APBWrittenWorkSCAR.pdf where A=interview number and B=participant number).
2. Copy videos to external Hard Drive –
    - Airdrop videos from iPhone and iPad to computer
    - Copy videos to external hard drive Scar Study\Original Videos\ named IV#P#Close and IV#P#Far
    - Delete videos from computer, iPhone and iPad
3. Use iMovie to create Audio and Both Views video
    - Open iMovie and create a new project, import the two movies (the close one first).
    - Create an Audio File –
        i. Assuming the iPad/close video has better sound recording, drag it to the timeline below.
        ii. Save the audio to an .mp3 file by selecting the share icon in the top right and selecting audio only (.mp3 if fine). Save file to external

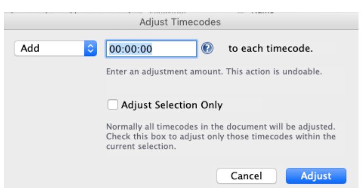
hard drive in Audio folder named IV#APBAudioSCAR where A=interview number and B=participant number.

- Create the both (2-video) file –
    i. Drag both movies to the timeline below (close view first).
    ii. Set Display to overlap and picture in picture (select top video in workspace to reveal overlay icon  on the top left of video display. Click on icon then choose picture in picture and resize far view to one of the corners.
    iii. Look for a similar spot on both videos to align them like the clap (spike in audios and) or lights turned off (darkness on videos) and use the mark function to indicate the same spot in each video and then the trim to playhead function (select spot in video and right click) and/or move one of the video feeds to align (the marked spots will show a line when they match). Check by playing with both audio observing no echo and observing picture in a picture screen video for alignment of actions.
    iv. Save the video to an .mp4 file by selecting the share icon in the top right and selecting video (fast/lowest quality is fine… if something is needed as higher quality it can be re-created later). Save file to external hard drive in Combined Video folder named IV#APBAudioSCAR where A=interview number and B=participant number.

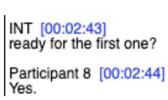
4. Create original otter.ai file – Upload the audio file to otter.ai . When the file is ready/processed, identify the speakers as Interviewer and P#-Name and save the file by clicking on the 3 vertical dots on the upper right and choosing “Export Text”. Save file to Transcripts folder with the name Otter.ai assigns (IV#APBAudioSCAR_otter.ai.txt where A=interview number and B=participant number).
5. Edit Transcript file – Using Otter.ai online editor, listen to file and make corrections. Save file as in 3 above, but add -edited. (IV#APBAudioSCAR_otter.ai-edited.txt)
6. Create Transcript file that can be uploaded to ATLAS.ti using insqribe. Copy and paste text from otter.ai-edited file into insqribe, then select Transcript>>Transcript Settings>> and check both “Omit frames” and “Recognize Unbracketed Settings”



Then select Transcript>>Adjust Timecodes>>



and select Adjust. The transcript should now have time stamps with square brackets and hours:munites:seconds)



Next select all (ctrl-A) and paste into a Word document and save as IV#APBTranscriptSCAR.doc. This is now a transcript file that ATLAS.ti can read and link to the video file.

APPENDIX B

INTERVIEW #1 TASKS AND ANSWERS

GOAL: Verify reasoning necessary for working with contingency tables and categorical association, which includes: (1) proportional reasoning, (2) basic probabilistic reasoning which includes identifying and redefining the sample space in context, and (3) the understanding of the structure of a contingency table both numerically (rows add across, columns add down, the total is both the marginal row total and the marginal column total) and contextually (each joint frequency has two different attributes, one for each variable and the structure and the marginal row frequencies represent the same observations in aggregate as the marginal column frequencies).

(1) Jessie is in charge of ordering pizzas for a school event. Each pizza has 8 slices and a teacher suggests on planning for 2 slices per person. If there are 47 people that are expected at the event, how may pizzas should Jessie order?

*A: 12; Since there are 8 slices per pizza and we are planning for 2 slices per person, we will need to order a pizza for every 4 people (4 people x 2 slices per person = 8 slices = 1 pizza). If there are 47 people, breaking them into groups of 4 will let us know how many pizza's to order (47 ÷ 4 = 11.75). We will have to round up, since there will be one group of 3 people and they still need pizza. OR If we need 2 slices per person and we have 47 people we will need 94 slices (2 * 47) and since pizzas have 8 slices each that is 11.75 pizzas (94 ÷ 8). OR The numbers of pizzas needed if each person had one slice is 5.875 pizzas (47÷ 8 = 5 7/8) and if each person had two slices, then we need double that or 11.75 pizzas (2 * 5.875)*

(2) The tallest student in the class reports to be 76 inches tall and the shortest student in the class reports to be 57 inches tall. A single book of unknown length is used to measure the tallest student who measures 8 books tall. Assuming the measurement is correct and the reported heights are correct, how many books lengths will the height of the shortest student be?

*A: 6; the book is 9 ½ or 19/2 inches tall (76 ÷ 8=9.5) and it will take 6 books to measure someone 57 inches (57 ÷ 19/2 = 6) OR we can say 76 is to 8 as 57 is to x and solve for x 76/8=57/x, x=57*8/76 =57*4/38 = 57*2/19 =114/19=6*

(3) In another class the tallest student claims to be 74 inches and the shortest student claims to be 56 inches tall. They each use a different book to measure their height. The tallest student reports to be 8 books tall and the shortest student reports to be 6 books tall. Are the books they used the same length?

A: no; The taller student's book is 9 ¼ or 37/4 inches tall (74 ÷ 8=9.25, how many 8's go into 74) and the shorter student's book is 9 1/3 or 28/3 inches tall (56 ÷ 6=9.33, how many 6's go into 56).

(4) Black and white paint are mixed together to create a certain shade of grey. This "dove grey" shade requires five (5) parts black paint to seven (7) parts white paint, or in other words the ratio of black to white paint is 5:7 or 5/7. Consider the following mixtures and determine if they are this shade of dove grey or not:

(a) 15 parts black paint and 21 parts white paint *(yes)*
(b) 25 parts black paint and 49 parts white paint *(no)*

A: (a) yes, [15:21] ÷ 3 = 5:7,
(b) no, [25:49] ÷ 5 =]5:9 4/5] or [25:49] ÷ 7 = [3 4/7 :7] neither of which are [5:7],
OR compute ratio as a decimal using smallest divided by largest 5/7 = 0.714
(a) yes, 15/21=0.714
(b) no, 25/49= 0.510
OR compute ratio as a decimal using largest divided by smallest 7 ÷ 5 = 1.4
(a) yes, 21 ÷ 15=1.4
(b) no, 49 ÷ 25= 1.96

(5) Consider the shade of "dove grey" paint with five (5) parts black paint to seven (7) parts white paint (the ratio of black to white paint is 5:7 or 5/7).
If the total mixture of paint is 3 gallons, how many gallons of black paint are needed?

*A: 5/12 = 5/4 = 1 ¼ = 1.25, When the total mixture has 12 gallons, five of the gallons are black, so the black paint is 5/12of the total "dove grey" paint. When there are 3 total gallons, the black paint is 5/12 of the 3 gallons, so 5/12 * 3 = 15/12.*

(6) Ming and Taylor are freshmen in a class of 24 students where 18 are freshmen and 6 are sophomores. One student from the entire class will be randomly chosen to present their project first. What is the probability that either Ming or Taylor will be chosen to present first?

A: 2/24 = 1/12, The event is 2 students that can be chosen, and the sample space is the 24 students.

(7) Consider the same situation, where Ming and Taylor are freshmen in a class of 24 students where 18 are freshmen and 6 are sophomores. But now, all freshmen will present their projects before any sophomores and the first freshman will be chosen at random. What is the probability Ming or Taylor will be chosen to present first?

A: 2/18 = 1/9 = 0.111, The event is 2 students that can be chosen, and the sample space is now the 18 freshmen students.

(8) You are at your school's soccer game and your trying to determine the likelihood of a student in the stands wearing glasses. You count the number of students and note there are 36 people wearing glasses, and 108 people who are not wearing glasses. Based on this information, what is your best estimate of the probability of a soccer fan at your school wearing glasses?

A: :36/144 = 1/4, The event is the 36 people with glasses and the sample space is the total crowd of 144 people (108 + 36).

(9) In a survey, a group of teens were asked several questions about their health. Two of the questions were *"Do you use e-cigarettes or vape?"* and *"Do you usually cough when you lie down to rest or sleep?"*. The responses are summarized in the following contingency table. What do you notice?

		Coughing habits		**Row Totals**
		Coughs when lying down	Does NOT Cough when lying down	
E-cig / vape usage	Do NOT use e-cig's/vape	41	137	178
	Uses e-cig's/vape	140	23	163
Column Totals		181	160	341

A: Numbers add across and down, more teens surveyed do NOT vape, more teens surveyed cough when lying down, teens surveyed who vape tend to cough when lying down…

(10) In a survey, a group of teens were asked several questions about their health. Two of the questions were *"Do you use exercise more than an hour a day?"* and *"Do you eat a well-balanced diet?"*. The responses are summarized in the following contingency table. Complete the missing values.

		Eating habits		**Row Totals**
		Eats a well-balanced diet	Does NOT eat a well-balanced diet	
Exercise frequency	Less than 1 hour per day of exercise	44	83	

	More than 1 hour per day of exercise	73	138	
Column Totals				

A:

44	*83*	*127*
73	*138*	*211*
117	*221*	*338*

(11) In a survey, a group of teens were asked several questions about their health. Two of the questions were *"Do you get an annual check-up with your doctor?"* and *"Do you visit your dentist every 6 months?"*. The responses are summarized in the following contingency table. Complete the missing values.

		Dentist visit frequency		Row Totals
		Regularly sees dentist	Does NOT regularly see dentist	
Annual doctor visit?	Gets annual check-up	143		189
	Does NOT get annual check-up			145
Column Totals		210	124	334

A:

143	*46*	*189*
67	*78*	*145*
210	*124*	*334*

(12) Do these two contingency tables represent the same or different results from a survey of teens that asks, *"Do you sleep more than 7 hours per night?"* and *"Do you drink more than 2 caffeinated beverages per day?"*?

		Caffeinated beverage frequency		Row Totals
		More than 2 drinks per day	2 or fewer drinks per day	
Hours of sleep at night	7 or more hours per night	43	78	121
	Less than 7 hours per night	142	85	227
Column Totals		185	163	348

	Hours of sleep at night		Row Totals
	7 or more hours per night	Less than 7 hours per night	

Caffeinated beverage frequency	2 or fewer drinks per day	78	85	163
	More than 2 drinks per day	43	142	185
Column Totals		121	227	348

A: same, all joint frequencies and marginal frequencies are the same

13) Do these two contingency tables represent the same or different results from a survey of teens that asks, *"Do you have frequent headaches?"* and *"Do you stretch regularly?"*?

		Headache frequency		Row Totals
		Frequently have headaches	Seldom have headaches	
Stretch frequency	Stretch frequently	49	32	81
	Stretch seldom	23	44	67
Column Totals		72	76	148

		Stretch frequency		Row Totals
		Stretch frequently	Stretch seldom	
Headache frequency	Frequently have headaches	49	32	81
	Seldom have headaches	23	44	67
Column Totals		72	76	148

A: different, based on marginal values(frequent headaches are 72 vs. 81, seldom headaches are 76 vs. 67, stretch frequently is 81 vs. 72, stretch seldom is 67 vs. 76) or based on some joint frequencies (stretch frequently and seldom headaches is 32 vs. 23 and stretch seldom and frequent headaches is 23 vs. 32)

(14) Consider the shade of "dove grey" paint with five (5) parts black paint to seven (7) parts white paint (the ratio of black to white paint is 5:7 or 5/7).
Consider the following mixtures and determine if they are this shade of paint or not:

(a) 2/5 parts black paint and 5/7 parts white paint *(no)*
(b) 2/7 parts black paint and 2/5 parts white paint *(yes)*

*A: (a) no, [2/5:5/7] *35 = [14:25] or [2/5:5/7] *5 = 2:3*
*(b) yes, [2/7:2/5] *35 = [10:14] and [10:14] ÷2 = [5:7]*
OR compute ratio as a decimal using smallest divided by largest 5/7 = 0.714
(a) no 2/5 ÷ 5/7 = 14/25 = 0.560
(b) yes 2/7 ÷ 2/5 = 0.714
OR compute ratio as a decimal using smallest divided by largest 7 ÷ 5 = 1.4

(a) no 2/5 ÷ 5/7 = 25 ÷ 14 = 1.79
(b) yes 2/5 ÷ 2/7 = 1.4

(15) Consider the shade of "dove grey" paint with five (5) parts black paint to seven (7) parts white paint (the ratio of black to white paint is 5:7 or 5/7).
If the total mixture of paint is 3 gallons, how many gallons of white paint are needed?

A:21/12 = 7/4 = 1 ¾ = 1.75 , When the total mixture has 12 gallons, seven of the gallons are white, so the white paint is 7/12 *of the total "dove grey" paint. When there are 3 total gallons, the black paint is* 7/12 *of the 3 gallons, so 7/12 * 3 = 21/12* OR *Knowing there is a total of 36/12 and black (from above) is 15/12, then white is 36/12 – 15/12 = 21/12.*

(16) In a survey, a group of teens were asked several questions about their health. Two of the questions were *"Do you often feel stressed?"* and *"Do you eat a lot of sweets?".* The responses are summarized in the following contingency table. Complete the missing values.

		Sweet eating frequency		**Row Totals**
		Eats many sweets	Eats few sweets	
Stress frequency	Often feel stressed	107	32	
	Seldom feel stressed		141	
Column Totals		170		

A:

107	*32*	*139*
63	*141*	*204*
170	*173*	*243*

(17) Considering any given contingency table, if the following actions occurred, would the resulting table represent the same or different situation?

(a) Switch the order of the rows including only the description (not the numbers)
(b) Switch the order of the columns including only the description (not the numbers)
(c) Switch the order of the rows including the description and numbers
(d) Switch the order of the columns including the description and numbers
(e) Switch the rows to columns and the columns to rows including only the description (not the numbers)
(f) Switch the rows to columns and the columns to rows including the description and the numbers

A: (a) different, (b) different, (c) same, (d) same, (e) different, (f) same

(18) Farah and Jamie run at the same pace around a track. Farah started before Jamie and when the coach arrived Farah had run a certain number of laps, say "*f*" laps by the time Jamie had run another number of laps, say "*j*" laps. When the coach leaves, Farah had run "*x*" laps. How many laps ("*y*") had Jamie run when the coach leaves?

A: $y = x - (f - j)$, For example, $f = 6$, $j = 3$, so when coach arrives, Farah had run 3 more laps. If an hour elapsed and they both run 7 laps an hour, they Farah would have run $x=f + 7 = 13$ and Jamie would have run $y=j + 10$. Since they run at the same rate, Farah always has run $f – j = 3$ more laps than Jamie.

APPENDIX C

INTERVIEW #2 TASKS AND ANSWERS

GOAL: To reveal reasoning about (in)dependence with completed contingency tables. Rather than using the words Association or Independence, questions are posed using the context of the problem and using the words more likely, less likely, or equally likely. A series of completed contingency tables with progressively challenging aspects is used.

(1) Consider the information in the following table which resulted from summarizing a survey question asking: *"If you could choose one way to be guaranteed to spend the rest of your life, would you rather choose to be happy or rich?"*

		Life preference		Row Totals
		Happy	Rich	
Grade band	Middle school	37	37	74
	High school	37	37	74
Column Totals		74	74	148

Are middle school students more, less, or equally likely than high school students to choose to be happy as opposed to rich?

A: Equally Likely, Middle school students (37/74) are equally likely as High school students (37/74) to choose to be happy as opposed to rich. All joint frequencies are the same (37=37=37=37), so therefore all marginal frequencies are the same (74=74=74=74) and all relative frequencies are the same 37/74=37/74=37/74=37/74=1/2=0.5).

(2) Consider the information in the following table which resulted from summarizing a survey question asking: *"If you could choose one way to be guaranteed to spend the rest of your life, would you rather choose to be happy or healthy?"*.

		Life preference		Row Totals
		Happy	Healthy	
Grade band	Middle school	74	0	74
	High school	0	74	74

Column Totals		74	74	148

Are middle school students more, less, or equally likely than high school students to choose to be happy as opposed to healthy?

A: More Likely, 74/74 > 0/74, or 74>0 and 74=74, or all MS prefer Happy and all HS prefer Healthy. Middle school students (74/74) are more likely then High school students (0/74) to choose to be happy as opposed to healthy.

(3) Consider the information in the following table which resulted from summarizing a survey question asking: *"If you could choose one way to be guaranteed to spend the rest of your life, would you rather choose to be healthy or rich?"*

		Life preference		**Row Totals**
		Healthy	Rich	
Grade band	Middle school	0	74	74
	High school	74	0	74
Column Totals		74	74	148

Are middle school students more, less, or equally likely than high school students to choose to be healthy as opposed to rich?

A: Less Likely, 0/74 < 74/74, or 0<74 and 74=74, or no MS prefer Healthy and all HS prefer Healthy. Middle school students (0/74) are less likely then High school students (74/74) to choose to be healthy as opposed to rich.

(4) Consider the information in the following table which resulted from summarizing a survey question asking: *"Which type of music do you prefer, rap or rock?"*

		Music preference		**Row Totals**
		Rap	Rock	
Grade band	Middle school	27	47	74
	High school	27	47	74
Column Totals		54	94	148

Are middle school students more, less or equally likely than high school students to prefer to listen to rap as opposed to rock?

A: Equally Likely, 27/74 = 27/74, or 27=27 and 74=74, or there an equal number of MS and HS students and the same number prefer rap. Middle school students (27/74) are equally likely as High school students (27/74) to choose to prefer rap as opposed to rock.

(5) Consider the information in the following table which resulted from summarizing a survey question asking, "Which type of music do you prefer, country or pop?"

		Music preference		Row Totals
		Country	Pop	
Grade band	Middle school	75	19	94
	High school	43	11	54
Column Totals		118	30	148

Are middle school students more, less or equally likely than high school students to prefer to listen to country as opposed to pop music?

A: Equally Likely, 75/94 ≅ 43/54, or 0.7978 ≅ 0.7962, or there an equal proportion of MS and HS students who prefer Country. Middle school students (73/94) are equally likely as High school students (43/54) to choose to prefer Country as opposed to Pop.

(6) Consider the information in the following table which resulted from summarizing a survey question asking, "Which type of music do you prefer, country or rock?"

		Music preference		Row Totals
		Country	Rock	
Grade band	Middle school	42	32	74
	High school	27	47	74
Column Totals		69	79	148

Are middle school students more, less, or equally likely than high school students to prefer to listen to country as opposed to rock music?

A: More Likely, 42/74 > 27/74, or 0.5675> 0.3648, or 42>27 and 74=74, or there are more MS students who prefer Country and the same amount of MS and HS students (74). or the percentage of MS students who prefer Country (57%) is greater than the percentage of HS students to prefer Country (36%). Middle school students (42/74) are more likely than High school students (27/74) to choose to prefer Country as opposed to Rock.

(7) Consider the information in the following table which resulted from summarizing a survey question asking, "Which type of music do you prefer, Rap or Country?"

		Music preference		Row Totals
		Rap	Country	
Grade band	Middle school	49	32	81
	High school	23	44	67
Column Totals		72	76	148

Are middle school students more, less, or equally likely than high school students to prefer to listen to Rap as opposed to Country music?

A: More Likely, 49/81 > 23/67, or 0.6049> 0.3432, or more than half of MS students prefer Rap whereas less than half of HS students prefer country, or the percentage of MS students who prefer Rap (60%) is greater than the percentage of HS students to prefer Rap (34%). Middle school students (49/81) are more likely than High school students (23/67) to choose to prefer Rap as opposed to Country.

(8) Students in Ms. Harvey's class at Westside High School were talking about different types of flu vaccines, a traditional shot and a more recently developed nasal mist. They decided to gather data from students and teachers at their high school. The following information summarizes which type of flu vaccine was taken and whether or not the flu was avoided or not.

		Flu status		**Row Totals**
		Avoided the flu	Got sick with flu	
Type of flu vaccine	Nasal mist	37	18	55
	Shot	31	24	55
Column Totals		68	42	110

Are students and teachers at Westside High School who received the Flu Mist more, less, or equally likely than those who got the flu shot to avoid the flu rather than get sick with the flu?

A: More Likely, 37/55 > 31/55, or 0.6727> 0.5636, or 37>31 and 74=74, or there are more people who avoided the flu that took the nasal mist (37) as opposed to receiving the shot ((31) and the same amount of people got the nasal mist and the shot (55), or the percentage of people who avoided the flu was greater for those who received the nasal mist (67%) that for those who received the shot (56%). Those who receive the nasal mist (37/55) are more likely than those who received the shot (31/55) to avoid getting the flu.

(9) Medical researchers at The Medline Clinic are testing two new drugs they think might cure a particular disease. The following information summarizes the drug given to patients at The Medline Clinic and whether or not the patient was cured.

		Disease status		**Row Totals**
		Cured	Not Cured	
Drug type	Drug A	29	9	38
	Drug B	46	22	68
Column Totals		75	31	106

Are patients at The Medline Clinic who receive Drug A more, less, or equally likely than patients who receive Drug B to be cured as opposed to not cured?

A: More Likely, 29/38 > 46/68, or 0.7631> 0.6764, or the percentage of people who were cured was greater for those who received Drug A (76%) as opposed to those who received Drug B (68%). Those who receive Drug A (29/38) are more likely than those who receive Drug B (46/68) to be cured as opposed to not be cured.

(9.b.) Consider the same table of information

		Disease status		Row Totals
		Cured	Not Cured	
Drug type	Drug A	29	9	38
	Drug B	46	22	68
Column Totals		75	31	106

Are patients at The Medline Clinic who are cured more, less, or equally likely than patients who are not cured to have received Drug A as opposed to Drug B?

A: More Likely, 29/75 > 9/31, or 0.3866> 0.2903, or the percentage of people who took Drug A was greater for those who were Cured (39%) as opposed to those who were not cured (29%). Those who were cured (29/75) are more likely than those who were not cured (9/31) to have received Drug A as opposed to receive Drug B.

(10) The tallest and shortest student in a classroom at Eastside High School are both left-handed. The tallest student claims that left-handed people tend to be taller and the shortest student claims that left-handed people tend to be shorter, each having studies to back up their findings. The class decides on criteria to classify students currently taking a psychology class in the high school as short or tall and gathers data to include their handedness. The information below summarizes the data.

		Height		Row Totals
		Taller	Shorter	
Hand preference	Right	42	36	78
	Left	21	18	39
Column Totals		63	54	117

At Eastside High School, are current psychology students who are left-handed likely to be taller or shorter than right-handed students?

A: Equally Likely, 42/78 = 21/39, or 0.5384 = 0.5384, or 21 is half of 42 and 78 is half of 39, or if I doubled the numbers for left-handed students, they would be equal to the *numbers for right-handed students. There are an equal percentage of right-handed (54%) and left-handed students (54%) who are taller. The right-handed students (42/78) are equally likely as the left-handed students (21/39) to be taller as opposed to shorter.*

(10.b.) Consider the same table of information

		Height		Row Totals
		Taller	Shorter	
Hand preference	Right	42	36	78
	Left	21	18	39
Column Totals		63	54	117

For current psychology students at Eastside High School, in comparison with shorter students, are taller students more likely to be right- or left-handed?

Equally Likely, 42/63 = 36/54, or 0.6666 = 0.6666, or 36 is 6/7 of 42 and 18 is 6/7 of 21, or if I multiplied the numbers for shorter students by 7/6, they would be equal to the numbers for right-handed students. There are an equal percentage of Taller (67%) and Shorter (67%) students who are right-handed. The Taller students (42/63) are equally likely as the Shorter students (36/54) to be taller as opposed to shorter.

(11) Breakfast cereal can be classified as kids' cereal or adults' cereal based on nutritional value such as grams of sugar and protein per serving. A class of high school students in the Evansville area visit local grocery stores to determine what shelf each kind of cereal is on, upper or lower. The following table summarizes their findings.

		Cereal location		Row Totals
		Upper shelf	Lower shelf	
Cereal type	Kids' cereal	62	47	109
	Adults' cereal	33	17	50
Column Totals		95	64	159

Based on this information, is a box of cereal in the Evansville area more likely to be on the upper or lower shelf because it is a kids' cereal as opposed to an adults' cereal?

Row conditional

57%	43%	100%
66%	34%	100%
60%	40%	100%

Column conditional

65%	73%	69%
35%	27%	31%
100%	100%	100%

A: A kids' cereal is more likely to be on the Upper Shelf, 62/109 > 47/109, or 57% > 43%.
Question may also be interpreted to ask about row conditional relative frequencies:
(1) Is it more (less or equally) likely to be on the Upper shelf because it is a Kids' cereal rather than an Adults' cereal? It is less likely to be on the Upper shelf because it is a Kids' cereal (62/109=57%) as opposed to an Adults' cereal (33/50=66%).
(2) Is it more (less or equally) likely to be on the Lower shelf because it is Kids' cereal rather than an Adults' cereal? It is more likely to be on the Lower shelf because it is a Kids' cereal (47/109=43%) as opposed to an Adults' cereal (17/50=34%).

(11.b.) Consider the same table of information

		Cereal location		**Row Totals**
		Upper shelf	Lower shelf	
Cereal type	Kids' cereal	62	47	109
	Adults' cereal	33	17	50
Column Totals		95	64	159

Based on this information, is a box of cereal in the Evansville area more likely to be a kids' cereal or an adults' cereal because it is on the upper shelf as opposed to the lower shelf?

A: More likely to be a Kids' cereal, 62/95 < 33/95 or 65%>35%.
Question may also be interpreted to ask about column conditional relative frequencies:
(1) Is it more (less or equally) likely to be a Kids' cereal because it is on the Upper shelf rather than the Lower shelf? It is Less likely to be a Kids' cereal because it is on the Upper shelf (62/95=65%) as opposed to the Lower shelf (47/64=73%).
(2) Is it more (less or equally) likely to be an Adults' cereal because it is on the Upper shelf rather than the Lower shelf? It is more likely to be an Adults' cereal because it is on the Upper shelf (33/95=35%) as opposed to the Lower shelf (17/64=27%).

APPENDIX D

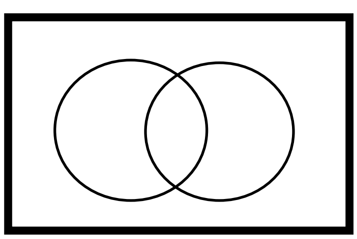
INTERVIEW #3 TASKS AND ANSWERS

GOAL: To assess the ability to create a valid Venn diagram from a contingency table and instruct how to create a mosaic plot. Also, to reveal reasoning about (in)dependence with completed contingency tables when an associated mosaic plot is provided. Rather than using the words Association or Independence, questions are posed using the context of the problem and using the words more likely, less likely, or equally likely. A series of completed contingency tables with progressively challenging aspects is used.

(Introductory Task) A survey was given to two different groups of students, middle and high school students. One of the questions asked about whether they spend more time using social media or playing video games. Here is a contingency table that summarizes the results:

Time spent		Social media	Video games	**Row Totals**
School grade	Middle School	10	40	50
	High School	75	75	150
Column Totals		85	115	200

(a) Complete a corresponding Venn diagram for the information in the contingency table:



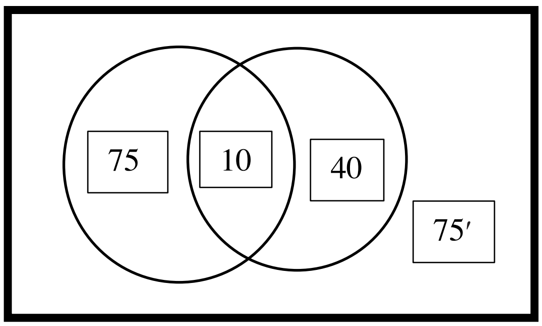
A*: The two circles can be labeled in the following ways in either order:*

1. *Middle School and Social Media (10 in the intersection)*
2. *Middle School and Video Games (40 in the intersection)*
3. *High School and Social Media (75 in the intersection)*

4. High School and Video Games (75 in the intersection)

(a.1) Complete the numbered corresponding Venn diagram for the information in the contingency table:

Time spent		Social media	Video games	**Row Totals**
School grade	Middle School	10	40	50
	High School	75	75′	150
Column Totals		85	115	200



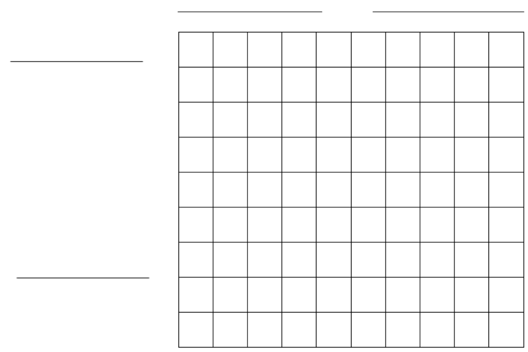
A: The first circle should be labeled Social Media and the second circle should be labeled Video Games

(a) A mosaic plot is another way to represent information in a contingency table. Since you may not be familiar with this type of graphical display, I will guide you through creating one for this first problem.

A survey was given to two different groups of students, middle and high school students. One of the questions asked about whether they spend more time using social media or playing video games. Here is a contingency table that summarizes the results:

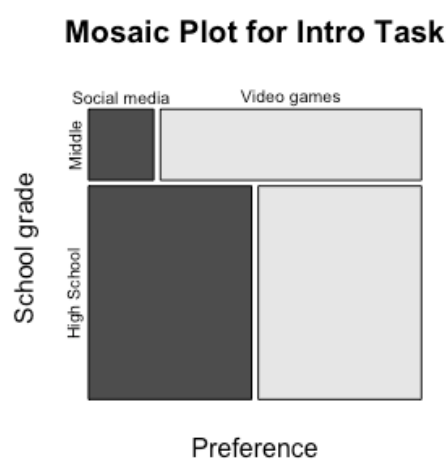
Time spent		Social media	Video games	**Row Totals**
School grade	Middle School	10	40	50
	High School	75	75	150
Column Totals		85	115	200

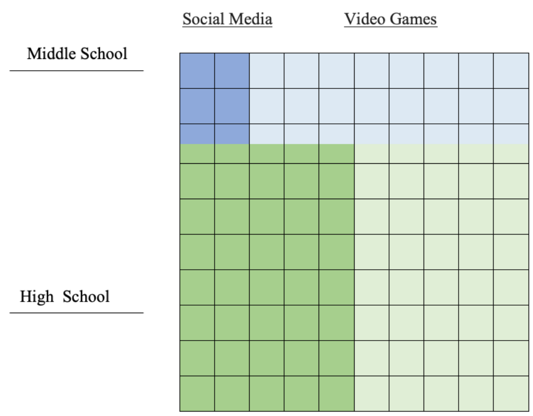
Here is a unit square with a 10 x 10 grid that we will use.



Here are some written instructions:

1. The first step is to divide the unit square using a horizontal line based on the row marginal frequencies, so in this case we will want to divide it based on the number of middle and high school students. Be sure to label these on the left side of your graph.
2. Next we consider each of these groups of middle and high school students separately and divide them with a vertical line based on the conditional frequencies. So in this case we will first consider the middle school students and divide the top row based on what proportion prefer social media and what proportion prefer video games. Be sure to label these on the top of your graph.
3. Similarly, partition the high school students.
4. Finally we will shade the graph. You can pick a color or pattern for each of the column variable categories, so in this instance one color or pattern for Social Media and another for Video Games.

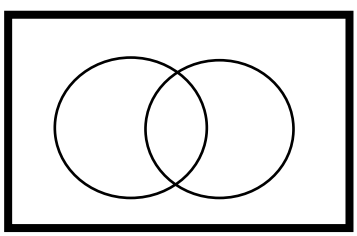




(1) Consider the information in the following table which resulted from summarizing a survey question asking: *"If you could choose one way to be guaranteed to spend the rest of your life, would you rather choose to be happy or rich?"*

		Life preference		**Row Totals**
		Happy	Rich	
Grade band	Middle school	37	37	74
	High school	37	37	74
Column Totals		74	74	148

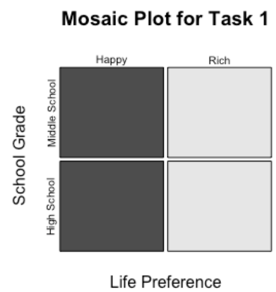
(a) Complete a corresponding Venn diagram for the information in the contingency table:



A: The two circles can be labeled in the following ways in either order:

1. *Middle School and Happy (37 in each area)*
2. *Middle School and Rich (37 in each area)*
3. *High School and Happy (37 in each area)*
4. *High School and Rich (37 in each area)*

(b) Complete a corresponding mosaic plot for the information in the contingency table:



(c) Are Middle school students more, less, or equally likely than High school students to choose to be Happy as opposed to Rich?

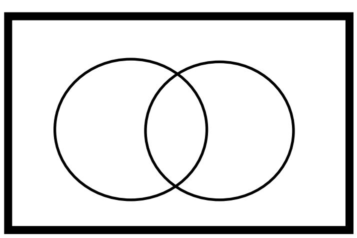
A: Equally likely

(2) Consider the information in the following table which resulted from summarizing a survey question asking: *"If you could choose one way to be guaranteed to spend the rest of your life, would you rather choose to be happy or healthy?"*.

		Life preference		**Row Totals**
		Happy	Healthy	
Grade band	Middle school	74	0	74
	High school	0	74	74
Column Totals		74	74	148

Are Middle school students more, less, or equally likely than High school students to choose to be Happy as opposed to Healthy?

(a) Complete a corresponding Venn diagram for the information in the contingency table:

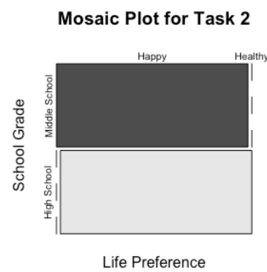


A: The two circles can be labeled in the following ways in either order:

1. *Middle School and Happy (74 in the intersection and outside, 0 in the circles)*
2. *Middle School and Healthy (0 in the intersection and outside,74 in the circles)*

3. *High School and Happy (74 in the intersection and outside, 0 in the circles)*
4. *High School and Healthy (0 in the intersection and outside, 74 in the circles)*

(b) Complete a corresponding mosaic plot for the information in the contingency table:



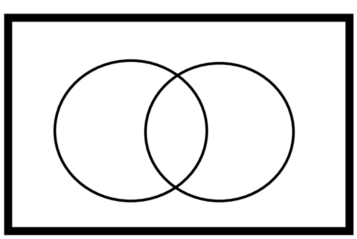
(c) Are Middle school students more, less, or equally likely than High school students to choose to be Healthy as opposed to Happy?

A: Less likely

(3) Consider the information in the following table which resulted from summarizing a survey question asking: *"If you could choose one way to be guaranteed to spend the rest of your life, would you rather choose to be healthy or rich?"*

		Life preference		**Row Totals**
		Healthy	Rich	
Grade band	Middle school	0	74	74
	High school	74	0	74
Column Totals		74	74	148

(d) Complete a corresponding Venn diagram for the information in the contingency table:

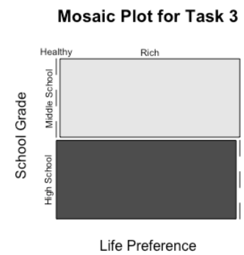


A: The two circles can be labeled in the following ways in either order:

5. *Middle School and Healthy (0 in the intersection and outside, 74 in the circles)*

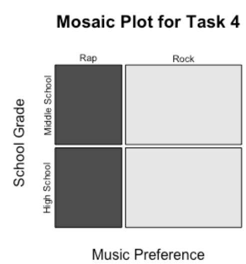
6. *Middle School and Rich (74 in the intersection and outside,0 in the circles)*
7. *High School and Healthy (0 in the intersection and outside, 74 in the circles)*
8. *High School and Rich (74 in the intersection and outside, 0 in the circles)*

(a) Complete a corresponding mosaic plot for the information in the contingency table:



(b) Are Middle school students more, less, or equally likely than High school students to choose to be Healthy as opposed to Rich?
A: Less Likely

(4) Consider the information in the following table which resulted from summarizing a survey question asking: *"Which type of music do you prefer, rap or rock?"*



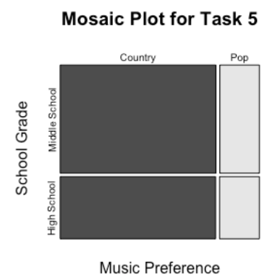
		Music preference		**Row Totals**
		Rap	Rock	
Grade band	Middle school	27	47	74
	High school	27	47	74
Column Totals		54	94	148

Are middle school students more, less, or equally likely than high school students to prefer to listen to rap as opposed to rock?

A: Equally Likely, 27/74 = 27/74, or 27=27 and 74=74, or there an equal number of MS and HS students and the same number prefer rap. Middle school students (27/74) are equally likely as High school students (27/74) to choose to prefer rap as opposed to rock.

Or, using the mosaic plot, the areas are the same or the horizontal length of the rectangles are the same.

(5) Consider the information in the following table which resulted from summarizing a survey question asking, "Which type of music do you prefer, country or pop?"



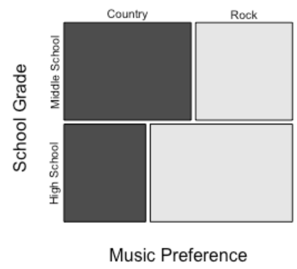
		Music preference		Row Totals
		Country	Pop	
Grade band	Middle school	75	19	94
	High school	43	11	54
Column Totals		118	30	148

Are middle school students more, less, or equally likely than high school students to prefer to listen to country as opposed to pop music?

A: Equally Likely, 75/94 = 43/54, or 0.7978 0.7962 27=27 and 74=74, or there an equal number of MS and HS students and the same number prefer rap. Middle school students (27/74) are equally likely as High school students (27/74) to choose to prefer Country as opposed to Pop.
Or, using the mosaic plot, the areas for the Country parts are in the same ratio to the Pop parts or the area for the country parts are the same fraction of the total area for the entire area for the middle or high school, or the horizontal length of the rectangles for country are the same, or there is a vertical line that splits the mosaic plot.

(6) Consider the information in the following table which resulted from summarizing a survey question asking, "Which type of music do you prefer, country or rock?"

Mosaic Plot for Task 6



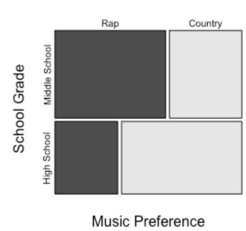
		Music preference		Row Totals
		Country	Rock	
Grade band	Middle school	42	32	74
	High school	27	47	74
Column Totals		69	79	148

Are middle school students more, less, or equally likely than high school students to prefer to listen to country as opposed to rock music?

A: More Likely, 42/74 > 27/74, or 0.5675> 0.3648, or 42>27 and 74=74, or there are more MS students who prefer Country and the same amount of MS and HS students (74). or the percentage of MS students who prefer Country (57%) is greater than the percentage of HS students to prefer Country (36%). Middle school students (42/74) are more likely than High school students (27/74) to choose to prefer Country as opposed to Rock.
Or, using the mosaic plot, the area (or horizontal length) of the middle school and Country is greater than the area (or horizontal length) of the high school and country. Could also consider ratio of area/length to Rock or total

(7) Consider the information in the following table which resulted from summarizing a survey question asking, "Which type of music do you prefer, rap or country?"

Mosaic Plot for Task 7



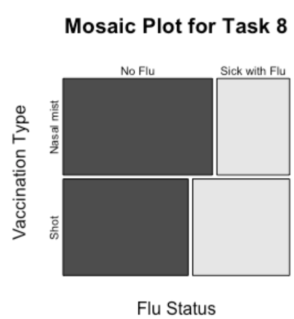
	Music preference	

		Rap	Country	Row Totals
Grade band	Middle school	49	32	81
	High school	23	44	67
Column Totals		72	76	148

Are middle school students more, less, or equally likely than high school students to prefer to listen to rap as opposed to country music?

A: More Likely, 49/81 > 23/67, or 0.6049> 0.3432, or more than half of MS students prefer Rap whereas less than half of HS students prefer country, or the percentage of MS students who prefer Rap (60%) is greater than the percentage of HS students to prefer Rap (34%). Middle school students (49/81) are more likely than High school students (23/67) to choose to prefer Rap as opposed to Country.
Or, using the mosaic plot, the area (or horizontal length) of the middle school and Rap is greater than the area (or horizontal length) of the high school and Rap. Could also consider ratio of area/length to Country or total.

(8) Students in Ms. Harvey's class at Westside High School were talking about different types of flu vaccines, a traditional shot and a more recently developed nasal mist. They decided to gather data from students and teachers at their high school. The following information summarizes which type of flu vaccine was taken and whether or not the flu was avoided or not.

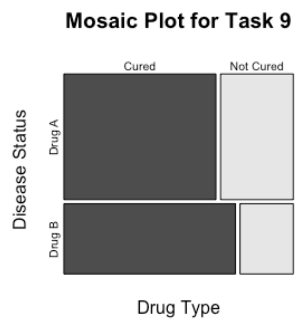


		Flu status		Row Totals
		Avoided the flu	Got sick with flu	
Type of flu vaccine	Nasal mist	37	18	55
	Shot	31	24	55
Column Totals		68	42	110

Are students and teachers at Westside High School who received the Flu Mist more, less, or equally likely than those who got the flu shot to avoid the flu rather than get sick with the flu?

A: More Likely, 37/55 > 31/55, or 0.6727> 0.5636, or 37>31 and 74=74, or there are more people who avoided the flu that took the nasal mist (37) as opposed to receiving the shot ((31) and the same amount of people got the nasal mist and the shot (55), or the percentage of people who avoided the flu was greater for those who received the nasal mist (67%) that for those who received the shot (56%). Those who receive the nasal mist (37/55) are more likely than those who received the shot (31/55) to avoid getting the flu. Or, using the mosaic plot, the area (or horizontal length) of the Nasal mist and Avoided the flu is greater than the area (or horizontal length) of the Shot and Avoided the flu. Could also consider ratio of area/length to Got sick with the flu or total.

(9) Medical researchers at The Medline Clinic are testing two new drugs they think might cure a particular disease. The following information summarizes the drug given to patients at The Medline Clinic and whether or not the patient was cured.

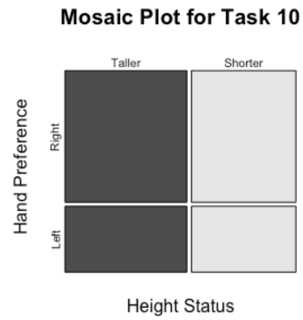


		Disease status		Row Totals
		Cured	Not Cured	
Drug type	Drug A	46	22	68
	Drug B	29	9	38
Column Totals		75	31	106

Are patients at The Medline Clinic who receive Drug A more, less, or equally likely than patients who receive Drug B to be cured as opposed to not cured?

A: Less Likely, 46/68< 29/38 or 0.6764 < 0.7631, or the percentage of people who were cured was less for those who received Drug A (68%) as opposed to those who received Drug B (76%). Those who receive Drug A (46/68) are less likely than those who receive Drug B (29/38) to be cured as opposed to not be cured.
Or, using the mosaic plot, the horizontal length of the Cured and Drug A is less than the horizontal length of the Cured and Drug B. Could also consider ratio of area/length to Not Cured or total.

(10) The tallest and shortest student in a classroom at Eastside High School are both left-handed. The tallest student claims that left-handed people tend to be taller and the shortest student claims that left-handed people tend to be shorter, each having studies to back up their findings. The class decides on criteria to classify students currently taking a psychology in the high school as short or tall and gathers data to include their handedness. The information below summarizes the data.

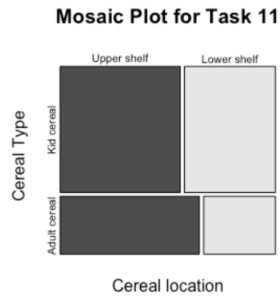


		Height		Row Totals
		Taller	Shorter	
Hand preference	Right	42	36	78
	Left	21	18	39
Column Totals		63	54	117

At Eastside High School, are current psychology students who are Left-handed likely to be Taller or Shorter than Right-handed students?

A: Equally Likely, 42/78 = 21/39, or 0.5384 = 0.5384, or 21 is half of 42 and 78 is half of 39, or if I doubled the numbers for left-handed students, they would be equal to the numbers for right-handed students. There are an equal percentage of right-handed (54%) and left-handed students (54%) who are taller. The right-handed students (42/78) are equally likely as the left-handed students (21/39) to be taller as opposed to shorter. Or, using the mosaic plot, the areas for the Taller parts are in the same ratio to the Shorter parts or the area for the Taller parts are the same fraction of the total area for the entire area for the Right- or Left-handed, or the horizontal length of the rectangles for Taller parts are the same, or there is a vertical line that splits the mosaic plot.

(11) Breakfast cereal can be classified as kids' cereal or adults' cereal based on nutritional value such as grams of sugar and protein per serving. A class of high school students in the Evansville area visit local grocery stores to determine what shelf each kind of cereal is on, upper or lower. The following table summarizes their findings.



		Cereal location		Row Totals
		Upper shelf	Lower shelf	
Cereal type	Kids' cereal	62	47	109
	Adults' cereal	33	17	50
Column Totals		95	64	159

Based on this information, is a box of cereal in the Evansville area more likely to be on the upper or lower shelf because it is a kids' cereal as opposed to an adults' cereal?

A: A Kids' cereal is more likely to be on the Upper Shelf, 62/109 > 47/109, or 57% > 43%.
Question may also be interpreted to ask about row conditional relative frequencies:
(1) Is it more (less or equally) likely to be on the Upper shelf because it is a Kids' cereal rather than an Adults' cereal? It is less likely to be on the Upper shelf because it is a Kids' cereal (62/109=57%) as opposed to an Adults' cereal (33/50=66%).
(2) Is it more (less or equally) likely to be on the Lower shelf because it is Kids' cereal rather than Adults' cereal? It is more likely to be on the Lower shelf because it is Kids' cereal (47/109=43%) as opposed to Adults' cereal (17/50=34%).
Using the mosaic plot,
(1) Less likely since the horizontal length of the Upper shelf and Kids' cereal rectangle is less than the horizontal length of the Upper shelf and Adults' cereal rectangle.
(2) More likely since the horizontal length of the Lower shelf and Kids' cereal rectangle is more than the horizontal length of the Lower shelf and Adults' cereal rectangle.

APPENDIX E

INTERVIEW #4 TASKS AND ANSWERS

GOAL: Assess students reasoning with incomplete contingency tables. Participants were given conditions of (in)dependence using the context of the problem and using the words less, equally, or more likely than as opposed to using the words association or independence.

(1a) Consider a contingency table which resulted from summarizing a survey question asking children and adults: *"If you could choose one between having a dog and a cat as a pet, which one would you choose?"*
Considering the total number in the contingency table below, what is a possible set of numbers that would cause you to conclude that Adults are MORE likely than children to prefer a cat rather than a dog?

		Pet preference		Row Totals
		Dog	Cat	
Age Category	Adults			
	Children			
Column Totals				168

A: answers may vary

0	*84*	*84*
84	*0*	*84*
84	*84*	*168*

42	*42*	*84*
80	*4*	*84*
122	*46*	*168*

4	*80*	*84*
42	*42*	*84*
46	*122*	*168*

(1b) Consider the same contingency table which resulted from summarizing a survey question asking children and adults: *"If you could choose one between having a dog and a cat as a pet, which one would you choose?"*

Considering the total number in the contingency table below, what is a possible set of numbers that would cause you to conclude that Adults are LESS likely than children to prefer a cat rather than a dog?

		Pet preference		Row Totals
		Dog	Cat	
	Adults			

Age Category	Children				
Column Totals					168

A: answers may vary

84	*0*	*84*
0	*84*	*84*
84	*84*	*168*

42	*42*	*84*
4	*80*	*84*
46	*122*	*168*

80	*4*	*84*
42	*42*	*84*
122	*46*	*168*

(1c) Consider a contingency table which resulted from summarizing a survey question asking children and adults: *"If you could choose one between having a dog and a cat as a pet, which one would you choose?"*

Considering the total number in the contingency table below, what is a possible set of numbers that would cause you to conclude that Adults are equally likely as compared to children to prefer a cat rather than a dog?

		Pet preference		**Row Totals**
		Dog	Cat	
Age Category	Adults			
	Children			
Column Totals				168

A: answers may vary

42	*42*	*84*
42	*42*	*84*
84	*84*	*168*

20	*64*	*84*
20	*64*	*84*
40	*128*	*168*

4	*80*	*84*
42	*42*	*84*
122	*48*	*168*

(2a) Students in Ms. Franklin's class at Northside High School were interested in finding out if there was a relationship between exercise frequency and sleeping problems. They decided to gather data from students and teachers at their high school.

Considering the total numbers in the contingency table below, what is a possible set of numbers that would cause you to conclude that the teachers and students at Northside Highschool who exercise frequently are LESS likely than those who do not exercise frequently to have sleeping problems as opposed to not have sleeping problems?

		Sleep problems?		**Row Totals**
		Do NOT have sleeping problems	Have sleeping problems	
Exercise Frequency	Exercise Frequently			86

	Do NOT Exercise Frequently			160
Column Totals				246

A: answers may vary

86	*0*	*86*
0	*160*	*160*
86	*160*	*246*

43	*43*	*86*
60	*100*	*160*
103	*143*	*246*

80	*6*	*86*
80	*80*	*160*
160	*86*	*246*

(2b) Students in Ms. Franklin's class at Northside High School were interested in finding out if there was a relationship between exercise frequency and sleeping problems. They decided to gather data from students and teachers at their high school.

Considering the total numbers in the contingency table below, what is a possible set of numbers that would cause you to conclude that the teachers and students at Northside Highschool who exercise frequently are EQUALLY likely when compared to those who do not exercise frequently to have sleeping problems as opposed to not have sleeping problems?

		Sleep problems?		**Row Totals**
		Do NOT have sleeping problems	Have sleeping problems	
Exercise Frequency	Exercise Frequently			86
	Do NOT Exercise Frequently			160
Column Totals				246

A: answers may vary

43	*43*	*86*
80	*80*	*160*
123	*123*	*246*

0	*86*	*86*
0	*160*	*160*
0	*246*	*246*

86	*0*	*86*
160	*0*	*160*
246	*0*	*246*

(2c) Students in Ms. Franklin's class at Northside High School were interested in finding out if there was a relationship between exercise frequency and sleeping problems. They decided to gather data from students and teachers at their high school.

Considering the total numbers in the contingency table below, what is a possible set of numbers that would cause you to conclude that the teachers and students at Northside Highschool who exercise frequently are MORE likely than those who do not exercise frequently to have sleeping problems as opposed to not have sleeping problems?

		Sleep problems?		Row Totals
		Do NOT have sleeping problems	Have sleeping problems	
Exercise Frequency	Exercise Frequently			86
	Do NOT Exercise Frequently			160
Column Totals				246

A: answers may vary

0	*86*	*86*
160	*0*	*160*
160	*86*	*246*

43	*43*	*86*
100	*60*	*160*
143	*103*	*246*

6	*80*	*86*
80	*80*	*160*
86	*160*	*246*

(3a) Medical researchers at The Wellness Clinic are testing two new drugs they think might cure a particular disease. The following information summarizes the drug given to patients at The Wellness Clinic and whether or not the patient was cured. Considering the total numbers in the contingency table below, what is a possible set of numbers that would cause you to conclude that Drug B is EQUALLY likely as Drug A to cure the disease for the patients at The Wellness Clinic?

		Disease status		Row Totals
		Cured	Still has disease	
Drug type	Drug A		12	86
	Drug B			160
Column Totals				246

A: answers may vary

0	*12*	*86*
160	*0*	*160*
160	*86*	*246*

43	*12*	*86*
100	*60*	*160*
143	*103*	*246*

6	*12*	*86*
80	*80*	*160*
86	*160*	*246*

(3b) Medical researchers at The Wellness Clinic are testing two new drugs they think might cure a particular disease. The following information summarizes the drug given to patients at The Wellness Clinic and whether or not the patient was cured.

Considering the total numbers in the contingency table below, what is a possible set of numbers that would cause you to conclude that Drug B is LESS likely than Drug A to cure the disease for the patients at The Wellness Clinic?

		Disease status		**Row Totals**
		Cured	Still has disease	
Drug type	Drug A		12	86
	Drug B			160
Column Totals				246

A: answers may vary

74	*12*	*86*
0	*160*	*160*
74	*172*	*246*

74	*12*	*86*
80	*80*	*160*
143	*103*	*246*

74	*12*	*86*
137	*23*	*160*
211	*35*	*246*

(3c) Medical researchers at The Wellness Clinic are testing two new drugs they think might cure a particular disease. The following information summarizes the drug given to patients at The Wellness Clinic and whether or not the patient was cured.

Considering the total numbers in the contingency table below, what is a possible set of numbers that would cause you to conclude that Drug B is MORE likely than Drug A to cure the disease for the patients at The Wellness Clinic?

		Disease status		**Row Totals**
		Cured	Still has disease	
Drug type	Drug A		12	86
	Drug B			160
Column Totals				246

A: answers may vary

74	*12*	*86*
160	*0*	*160*
234	*12*	*246*

74	*12*	*86*
150	*10*	*160*
224	*23*	*246*

74	*12*	*86*
138	*22*	*160*
212	*34*	*246*

(4a) Breakfast cereal can be either name brand or the store brand. A class of high school students in the Blairsville area visit local grocery stores to determine what shelf each kind of cereal is on, upper or lower.
Considering the total numbers in the contingency table below, what is a possible set of numbers that would cause you to conclude that in the Blairsville area, the Name-brand

cereal is EQUALLY likely when compared to the Store brand cereal to be on the Lower shelf?

		Shelf location		Row Totals
		Upper shelf	Lower shelf	
Cereal type	Name-brand cereal			70
	Store-brand cereal			118
Column Totals		102	86	188

A: Answer

38	*32*	*70*
64	*54*	*118*
102	*86*	*188*

(4b) Breakfast cereal can be either name brand or the store brand. A class of high school students in the Blairsville area visit local grocery stores to determine what shelf each kind of cereal is on, upper or lower.

Considering the total numbers in the contingency table below, what is a possible set of numbers that would cause you to conclude that in the Blairsville area, the Name-brand cereal is MORE likely than the Store-brand cereal to be on the Lower shelf?

		Shelf location		Row Totals
		Upper shelf	Lower shelf	
Cereal type	Name-brand cereal			70
	Store-brand cereal			118
Column Totals		102	86	188

A: answers may vary

35	*35*	*70*
67	*51*	*118*
102	*86*	*188*

0	*70*	*70*
102	*16*	*118*
102	*86*	*188*

2	*68*	*70*
100	*18*	*118*
102	*86*	*188*

(4c) Breakfast cereal can be either name brand or the store brand. A class of high school students in the Blairsville area visit local grocery stores to determine what shelf each kind of cereal is on, upper or lower.

Considering the total numbers in the contingency table below, what is a possible set of numbers that would cause you to conclude that in the Blairsville area, the Name-brand cereal is LESS likely than the Store-brand cereal to be on the Lower shelf?

		Shelf location		Row Totals
		Upper shelf	Lower shelf	
Cereal type	Name-brand cereal			70
	Store-brand cereal			118
Column Totals		102	86	188

A: answers may vary

43	*27*	*70*
59	*59*	*118*
102	*86*	*188*

70	*0*	*70*
32	*86*	*118*
102	*86*	*188*

60	*10*	*70*
40	*78*	*118*
102	*86*	*188*

(5a) Medical researchers at The Wellness Clinic are testing two new drugs they think might cure a particular disease. The following information summarizes the drug given to patients at The Wellness Clinic and whether or not the patient was cured.

Considering the total numbers in the contingency table below, what is a possible set of numbers that would cause you to conclude that the patients who were cured were EQUALLY likely as those who still had the disease to have gotten Drug C for the patients at The Wellness Clinic?

		Disease status		Row Totals
		Cured	Still has disease	
Drug type	Drug C		32	256
	Drug D			64
Column Totals				320

A:

224	*32*	*256*
56	*8*	*64*
280	*40*	*320*

(5b) Medical researchers at The Wellness Clinic are testing two new drugs they think might cure a particular disease. The following information summarizes the drug given to patients at The Wellness Clinic and whether or not the patient was cured.

Considering the total numbers in the contingency table below, what is a possible set of numbers that would cause you to conclude that patients who were cured were LESS likely those who still had the disease to have gotten Drug C for the patients at The Wellness Clinic?

		Disease status		Row Totals
		Cured	Still has disease	
Drug type	Drug C		32	256
	Drug D			64
Column Totals				320

A: answers may vary

224	*32*	*256*
0	*64*	*64*
224	*96*	*320*

224	*32*	*256*
32	*32*	*64*
256	*64*	*320*

224	*32*	*256*
55	*9*	*64*
279	*41*	*320*

(5c) Medical researchers at The Wellness Clinic are testing two new drugs they think might cure a particular disease. The following information summarizes the drug given to patients at The Wellness Clinic and whether or not the patient was cured.

Considering the total numbers in the contingency table below, what is a possible set of numbers that would cause you to conclude that patients who were cured were LESS likely those who still had the disease to have gotten Drug C for the patients at The Wellness Clinic?

		Disease status		Row Totals
		Cured	Still has disease	
Drug type	Drug C		32	256
	Drug D			64
Column Totals				320

A: answers may vary

224	*32*	*256*
64	*0*	*64*
288	*32*	*320*

224	*32*	*256*
62	*2*	*64*
286	*34*	*320*

224	*32*	*256*
57	*7*	*64*
281	*39*	*320*

APPENDIX F

INTERVIEW #5 TASKS AND ANSWERS

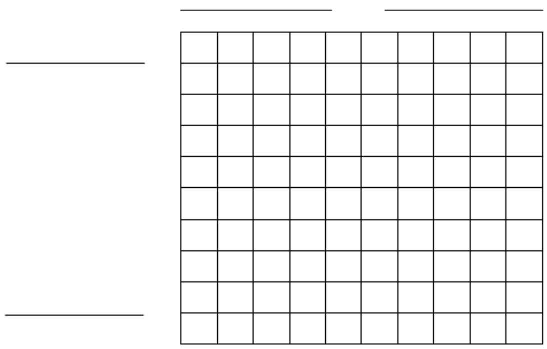
GOAL: Assess students ability to draw a representation of their choosing as well as their recollection of drawing a mosaic plot. Gain an understanding of participant's current understanding of some terms (association and independence; category and variable). Assess participants' reasoning with incomplete contingency tables and mosaic plots when both are provided. Participants were given conditions of (in)dependence using the context of the problem and using the words less, equally, or more likely than as opposed to using the words association or independence.

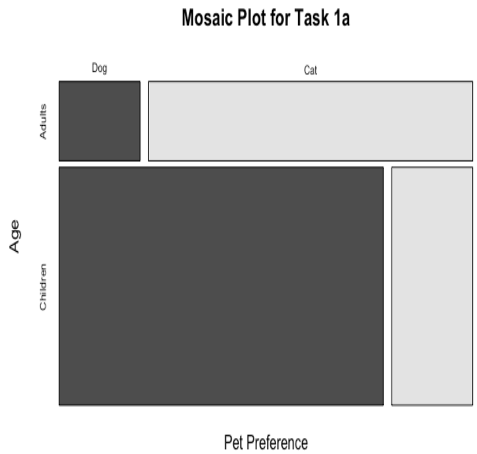
Preliminary questions:
What grade are you in?
How old are you?
What math class are you currently taking?
What science class are you currently taking?
What are your favorite subjects?
What do you like to do when you are not in school?
What would help you like math more?
What would help you improve your understanding of math?

(1a) Consider a contingency table which resulted from summarizing a survey question asking children and adults: *"If you could choose one option between having a dog and a cat as a pet, which one would you choose?"*
Considering the following result from a group of people which causes you to conclude that Adults are **MORE** likely than children to prefer a cat rather than a dog: What other representation might you draw to visually show this situation?

		Pet preference		Row Totals
		Dog	Cat	
Age Group	Adults	10	40	50
	Children	120	30	150
Column Totals		130	70	200

How would you draw a mosaic plot to visually show this situation?

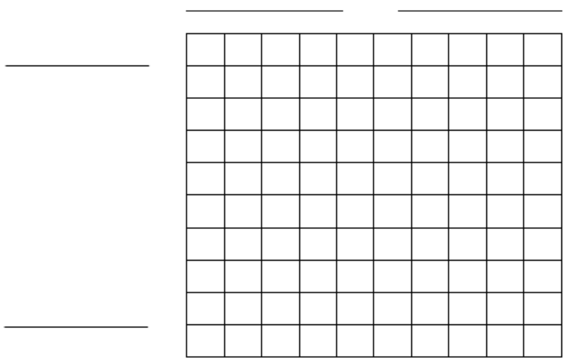


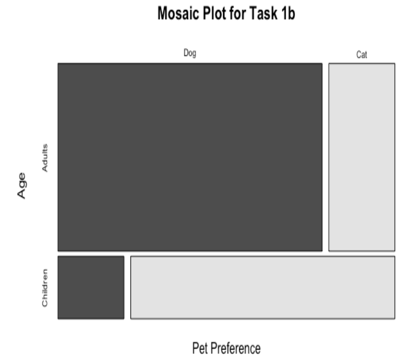


(1b) The same survey question is asked to a different group of children and adults: *"If you could choose one option between having a dog and a cat as a pet, which one would you choose?"*
Considering the following result from a different group of people which causes you to conclude that Adults are **<u>LESS</u>** likely than children to prefer a cat rather than a dog:

		Pet preference		**Row Totals**
		Dog	Cat	
Age Group	Adults	120	30	150
	Children	10	40	50
Column Totals		130	70	200

What do you think a mosaic plot might look like for this situation?

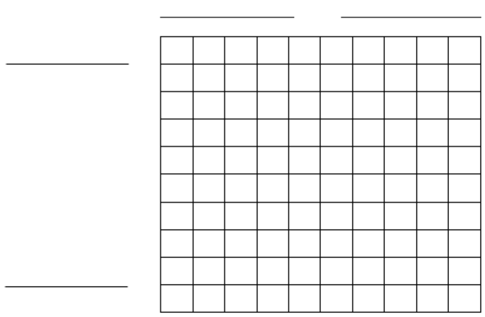


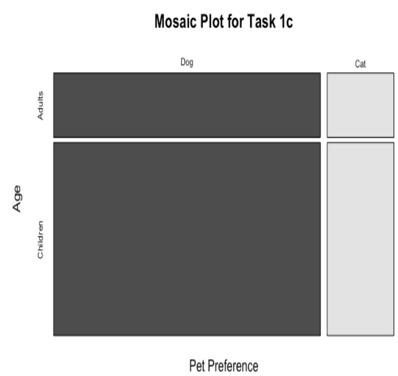


(1c) The same survey question is asked to a different group of children and adults: *"If you could choose one option between having a dog and a cat as a pet, which one would you choose?"*
Considering the following result from a different group of people which causes you to conclude that Adults are **EQUALLY** likely when compared to children to prefer a cat rather than a dog:

		Pet preference		**Row Totals**
		Dog	Cat	
Age Group	Adults	40	10	50
	Children	120	30	150
Column Totals		160	40	200

What do you think a mosaic plot might look like for this situation?





1. Statisticians use the terms <u>independent</u> and <u>association</u>. What do these terms mean to you?

<u>Independent:</u>

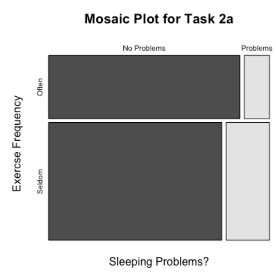
<u>Association:</u>

2b. Statisticians also use the terms <u>category</u> and <u>variable.</u> What do these terms mean to you?

<u>Category:</u>

<u>Variable:</u>

(2a) Students in Ms. Franklin's class at Northside High School were interested in finding out if there was a relationship between exercise frequency and sleeping problems. They decided to gather data from students and teachers at their high school.
Considering a the mosaic plot and the total numbers in the contingency table below, what is a possible set of numbers that would cause you to conclude that the teachers and students at Northside Highschool who exercise frequently are **<u>LESS</u>** likely than those who do not exercise frequently to have sleeping problems as opposed to not have sleeping problems?



		Sleep problems?		**Row Totals**
		Do NOT have sleeping problems	Have sleeping problems	
Exercise Frequency	Exercise Frequently			86
	Do NOT Exercise Frequently			160
Column Totals				246

A: answers may vary

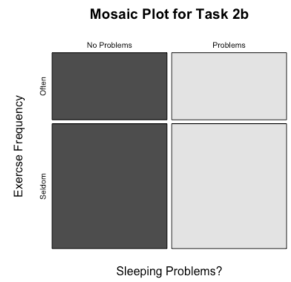
70	*16*	*86*
120	*40*	*160*
190	*56*	*246*

74	*12*	*86*
125	*35*	*160*
299	*47*	*246*

76	*10*	*86*
130	*30*	*160*
206	*40*	*246*

(2b) Consider Ms. Franklin's class and some different results of the survey.
Considering the mosaic plot and the total numbers in the contingency table below, what is a possible set of numbers that would cause you to conclude that the teachers and students at Northside Highschool who exercise frequently are **<u>EQUALLY</u>** likely when

compared to those who do not exercise frequently to have sleeping problems as opposed to not have sleeping problems?



		Sleep problems?		Row Totals
		Do NOT have sleeping problems	Have sleeping problems	
Exercise Frequency	Exercise Frequently			86
	Do NOT Exercise Frequently			160
Column Totals				246

A: answers may vary

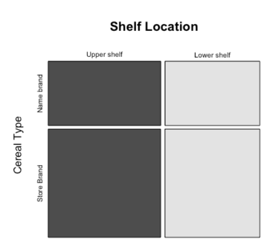
43	*43*	*86*
80	*80*	*160*
123	*123*	*246*

35	*51*	*86*
65	*95*	*160*
100	*146*	*246*

39	*47*	*86*
72	*88*	*160*
111	*135*	*246*

(3) Breakfast cereal can be either name brand or the store brand. A class of high school students in the Blairsville area visit local grocery stores to determine what shelf each kind of cereal is on, upper or lower.

Considering the mosaic plot and the total numbers in the contingency table below, what is a possible set of numbers that would cause you to conclude that in the Blairsville area, the Name-brand cereal is **EQUALLY** likely when compared to the Store-brand cereal to be on the Lower shelf?



		Shelf location		Row Totals
		Upper shelf	Lower shelf	
Cereal type	Name-brand cereal			70
	Store-brand cereal			118
Column Totals		102	86	188

A: Answer

38	*32*	*70*
64	*54*	*118*
102	*86*	*188*