

The Difficulty and Accessibility of Combinatorics Problems: Evidence from Large-scale
Assessments and Student Interviews

Fernando Amaral Carnauba

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2024

© 2024

Fernando Amaral Carnauba

All Rights Reserved

Abstract

The Difficulty and Accessibility of Combinatorics Problems:
Evidence From Large-scale Assessments and Student Interviews

Fernando Amaral Carnauba

The objective of this dissertation was to explore the paradoxical nature of Combinatorics as both a difficult and accessible domain in Mathematics, particularly for K-12 students. This paradox in Combinatorics' nature raised questions about how students interact with problems in this domain and the factors influencing their understanding and engagement with mathematics.

To investigate these aspects, the study utilized a mixed-methods approach. Quantitative data was derived from the Exame Nacional do Ensino Médio (ENEM), a large-scale nationwide assessment in Brazil. The analysis focused on 28 Combinatorics problems identified across 12 years of the exam, comparing them with non-Combinatorics problems. The study also involved qualitative methods, specifically task-based interviews with Brazilian students, primarily from disadvantaged school backgrounds. These interviews aimed to provide deeper insights into how students approach, understand, and engage with Combinatorics problems.

The findings revealed that while the combinatorial domain is notably accessible in the sense that it allows students with varied backgrounds to understand what problems ask, this accessibility does not necessarily translate into students consistently arriving at correct solutions. The study also found that achievement gaps between students of private and public schools in Brazil are smaller in Combinatorics than in other mathematical domains. Together, these findings point to Combinatorics as a domain that can contribute to issues of equity in mathematics teaching and learning. Furthermore, the research underscored the importance of

considering both the 'product' (correct answers) and 'process' (mathematical thinking) aspects in mathematics education, especially in contexts aiming to promote equitable learning opportunities.

Table of Contents

List of Figures	iii
List of Tables	v
Acknowledgments.....	vi
Chapter 1: Introduction.....	1
Need for the Study	1
Purpose.....	5
Procedures.....	5
Chapter II: Literature Review	10
The Domain of Combinatorics.....	10
Combinatorics: A Paradoxically Accessible and Difficult Domain?	13
Combinatorics as an Accessible Domain.....	18
Combinatorics as a Difficult Domain	27
Educational Inequalities in Brazil.....	34
Chapter III: Methodology	38
Quantitative Data and Research Questions 1 and 2	38
Qualitative Data and Research Question 3	55
Chapter IV: Results.....	67
Data overview	67
Research Question 1	68
Research Question 2	76

Research Question 3	80
Chapter V: Conclusion.....	123
Summary of Results	123
Discussion and Contributions	125
Limitations and Pathways for Research.....	131
References.....	133
Appendix A: Interview Protocol.....	141
Appendix B: Interview Problems.....	143
Appendix C: Additional Tables	156

List of Figures

Figure 2.1: Difficulty and Accessibility in Mathematics Problems.....	17
Figure 2.2: Schematic representation of hypotheses drawn from the literature	34
Figure 3.1: Example of ENEM problem classified as in the Combinatorics domain.....	40
Figure 3.2: Example of ENEM problem classified as not in the Combinatorics domain.....	41
Figure 3.3: Empirical procedure for the first part of Research Question 1 (difficulty)	44
Figure 3.4: Hypothetical maximum school achievement gaps for varied difficulty levels	47
Figure 3.5: Empirical procedure for the first part of Research Question 1 (accessibility)	48
Figure 3.6: Hypotheses tested to address Research Question 1	50
Figure 3.7: Empirical procedure for the first part of Research Question 2.....	54
Figure 3.8: Hypotheses tested to address Research Question 2.....	55
Figure 4.1: Difficulty level distributions of Combinatorics and non-Combinatorics problems in ENEM (2009-2020, grey test version).....	69
Figure 4.2: School and racial gap distributions of Combinatorics and non-Combinatorics problems in ENEM (2009-2020, grey test version).....	72
Figure 4.3: Coding themes used in the qualitative analysis.....	84
Figure 4.4: Problem 6, Airplane Seats	86
Figure 4.5: Problem 10, Toy Train	88
Figure 4.6: Problem 1, Detective Game.....	93
Figure 4.7: Problem 11, Motorbike Trip.....	94
Figure 4.8: Problem 7, Painting Circles.....	95
Figure 4.9: Problem 8, Scouts' Tent	96
Figure 4.10: Students' annotations in Problem 7, Painting Circles	99

Figure 4.11: Problem 12, Paths on Map	100
Figure 4.12: Students' annotations and quotes in Problem 12, Paths on Map.....	101
Figure 4.13: Problem 3, Bank Password.....	102
Figure 4.14: Problem 9, Volleyball Teams.....	103
Figure 4.15: Problem 5, Hiring Process.....	114
Figure 4.16: Problem 4, Renting Movies.....	117
Figure 5.1: Schematic representation of updated hypotheses	127

List of Tables

Table 3.1. Criteria for the first round of coding.....	63
Table 4.1. Combinatorics and non-Combinatorics items used in the study.....	67
Table 4.2. Descriptive statistics of subpopulation ENEM test-takers of the study.....	68
Table 4.3. Regressions of difficulty levels (dependent variable) on a Combinatorics dummy and covariates (ENEM 2009-2020, grey test version).....	70
Table 4.4: Regressions of private-public school achievement gaps (an indicator of accessibility, in the opposite direction) on a Combinatorics	74
Table 4.5: Regressions of black-white achievement gaps (an indicator of accessibility, in the opposite direction) on a Combinatorics dummy and covariates	74
Table 4.6: Quantifiable characteristics of Combinatorics and non-Combinatorics items used in the study	77
Table 4.7: New regression models including observable characteristics (length and notation)...	78
Table 4.8. Indicators of accessibility and difficulty in student interviews	82
Table 4.9. Pre-identified potentially challenging words and symbols.....	85
Table 4.10: Answer-based guessing	116

Acknowledgments

Writing this dissertation would not have been a worthwhile – or even possible – experience without the support of many people around me. I extend my heartfelt gratitude to my friends at TC and NYC. Murilo, Mine, Chen Chen, Danilo, Fernanda, and Renato, thank you for sharing the highs and lows of graduate school.

The Mathematics Education program at Teachers College provided an unparalleled environment during my Ph.D. journey. I am sincerely thankful for the enriching experiences in mathematics, mathematics education, and research methods. Dr. Blikstein, thank you for welcoming me to my TC journey and for demonstrating the collaborative and international nature of research. Dr. Karp, thank you for your dedication to supporting students and your truthful contributions to classroom discussions. Dr. Smith, your passion for collaborative mathematics has left a lasting impression, and I aspire to teach math as inspiringly as you do.

Dr. Wasserman, as my dissertation advisor, your inspiration through research, teaching, and guidance was invaluable. Your dedication and professionalism set a standard in our field, and I am deeply grateful for your patience, attentive listening, and thoughtful guidance during the dissertation writing process.

My sincere thanks to my friends at Instituto Canoa and the Stanford Graduate School of Education for their unwavering support. Andréa, César, Fran, Gui – our partnership over the years has been inspiring. To Alessandra, Anderson, Bilzã, Fábio, Fransueli, Kaique, Lolla, Pati, Rita, Stephani, Tati – I have learned immensely from each of you. Mila, your leadership is a constant source of learning, and I am blessed to have found a partner in education with you.

I am privileged to have had the opportunity to know and collaborate with some of my idols. Dr. Carnoy, your belief in Canoa's work and efforts to improve Brazilian education are

deeply appreciated. Dr. Lotan, thank you for inspiring and supporting not only me but an entire generation of Brazilian scholars. You are a lighthouse.

This dissertation would not have been possible without the encouragement and support of my loving family. Rui and Chica, thank you for always welcoming us to your place and making us feel at home. Mamãe, Papai, Dan, Gui, Vovó Lygia, thank you for always being there for me - I love you with all my heart. Julia and Alice, your love is my greatest blessing. I love you, and you are everything to me.

Chapter 1: Introduction

This chapter provides a brief overview of this dissertation study. First, it introduces the need for the study by discussing a puzzle within the mathematics education literature about the domain of Combinatorics. Second, it presents the study's objectives by outlining the research questions it addresses. The chapter concludes with a concise summary of the study's methodology.

Need for the Study

In recent decades, the mathematics education community has shown growing interest in Discrete Mathematics, with a particular focus on Combinatorics. Scholars underscored the significance of Combinatorics as a foundation for probability, computer science, and the development of students' problem-solving skills (DeBellis & Rosenstein, 2004; Kapur, 1970; Tucker, 2012). However, the literature is less clear about whether and how Combinatorics problems are accessible for K-12 students. Some studies suggest that Combinatorics is accessible to students with diverse backgrounds and varying levels of mathematical proficiency (DeBellis & Rosenstein, 2004; Goldin, 2004; Hurdle et al., 2016; Lockwood et al., 2020). Conversely, research has also documented varied difficulties students encounter when tackling Combinatorics problems, indicating that it may be a challenging domain to teach and learn (Batanero et al., 1997; Dubois, 1984; Hadar & Hadass, 1981). This contrasting evidence about how students interact with Combinatorics problems suggests that further investigation of this domain may provide insights into the factors influencing accessibility in mathematics education.

The most salient explanation associated with the claim that Combinatorics problems are accessible is that they do not require extensive familiarity with mathematics-specific vocabulary and notation (Lockwood & Gibson, 2016; Lockwood et al., 2020; Sandefur et al., 2022).

Students who had previously struggled with school mathematics can thus understand what Combinatorics problems ask, as these “tend to be accessible problems that require little mathematical background” (Sandefur et al., 2022, p.761).

Researchers also highlight that Combinatorics problems are generally difficult to solve. Kapur (1970) explains that Discrete Mathematics - the broader area of mathematics that encompasses Combinatorics - frequently does not allow for “powerful methods” to be applied in solving a wide range of problems (p.114). Thus, knowing formulas or pre-determined methods does not suffice for students to solve Combinatorics problems correctly. Accordingly, Batanero et al. (1997) found that the most common mistake incurred by high school students working on Combinatorics problems was the failed attempt to use an algorithm - the formula for combinations or arrangements - in a context to which it did not apply.

This dissertation addresses the seeming contradiction that Combinatorics is both difficult and accessible by distinguishing the concepts of *product* and *process* in students' interactions with mathematics problems. A product perspective emphasizes students' final answers, focusing on whether students can provide correct solutions to Combinatorics problems. It is under this product lens that the literature describes Combinatorics as a difficult domain - one in which students struggle to provide correct final answers. Research documenting these challenges, such as Batanero et al. (1997) and Hadar and Hadass (1981), reveals various mistakes students make when solving Combinatorics problems.

Differently, the claims that Combinatorics is an accessible domain usually emphasize the processes with which it allows students to engage. Interest in the process aspect of students' interactions with mathematics problems is prevalent in the mathematics education literature. Influential books and articles, such as Polya's (1954) and Schoenfeld's (1985) works on

problem-solving, and Cuoco et al.'s (1996) work on mathematics habits of mind, focus on understanding the processes involved in solving mathematical problems. They emphasize the importance of providing students with opportunities to think like mathematicians. The National Council of Teachers of Mathematics (NCTM) contributed to this interest by including process standards in their Principles and Standards publication (NCTM, 2000), which later influenced the Common Core State Standards for Mathematics' Mathematical Practices (National Governors Association Center for Best Practices, 2010).

The concept of mathematical processes offers insight into assertions regarding the accessibility of Combinatorics. The literature describes this domain as one in which students can understand problems and engage in solution attempts that spark valuable mathematical thinking. Students may engage in valuable mathematical processes while attempting a solution, even when they do not arrive at the correct answer. Combinatorics problems are described as accessible in the sense that they enable students with varied mathematical backgrounds to work toward a solution (English, 2005; Lockwood et al., 2020). If these solution attempts involve valuable mathematical processes, such as those defined by Cuoco et al. (1996) or NCTM (2000), students can benefit from engaging with Combinatorics problems even if they do not always reach the correct answers. Combinatorics is thus a domain described in the literature as offering extensive opportunities for engaging in significant mathematical reasoning (process) while also presenting challenges that many students find difficult to overcome when attempting to solve problems accurately (product). To disentangle the concepts of difficulty and accessibility, this dissertation uses the following definitions:

Difficulty: *The rate of incorrect responses provided by students in a mathematics problem.*

Accessibility: *The extent to which students understand what a mathematics problem asks and engage in mathematically valuable processes while attempting to solve it.*

Although previous research has documented students' interactions with Combinatorics problems in varied contexts, there are two reasons for further investigation. First, the current body of evidence does not warrant the conclusion that Combinatorics is more (or less) accessible and difficult to students *than other domains in mathematics*. Only empirical comparisons of the teaching and learning of Combinatorics with other domains, which still lack in the literature, can be conclusive on this matter. The importance of a direct comparison between Combinatorics and non-combinatorics problems cannot be overstated. Teaching and laboratory experiments in the Combinatorics literature are often designed and implemented by the author-researchers themselves. If researchers of Combinatorics are particularly fond of specific teaching approaches - for example, favoring problem-solving and student conceptual understanding over the memorization of algorithms - then the outcomes of such experiments might result from the teaching methods used rather than from particularities of Combinatorics problems. Considering the growing attention of mathematics education researchers toward Combinatorics in secondary schools and its potential contribution to achieving equity in mathematics learning (Lockwood et al., 2020), it is worthwhile to deepen our comprehension of this domain's unique characteristics relative to other mathematical domains.

Second, to the best of my knowledge, no study has explicitly examined the specific characteristics of Combinatorics problems that may be associated with accessibility (process) and difficulty (product). Although past research outside the Combinatorics literature, such as the works of Avalos et al. (2018) and Martiniello (2009), suggests that linguistic factors may

significantly influence students' opportunities to engage with mathematics problems, these findings have not been replicated in the context of Combinatorics. As mathematics educators often cite the straightforward language used in Combinatorics problems as a potential contributing factor to its accessibility, it is crucial to explore this connection through empirical investigation.

Purpose

This study investigates the determinants of and the relationship between difficulty and accessibility in Combinatorics problems. The following research questions are proposed:

- (1) How do Combinatorics problems compare to problems in other domains in terms of difficulty and accessibility?*
- (2) What (if any) are some characteristics of Combinatorics problems that are associated with their difficulty and accessibility levels?*
- (3) What are some of the ways in which students access problems in Combinatorics? What are some of the ways in which students fail to provide correct answers to problems in Combinatorics? What is the relationship between difficulty and accessibility in students' interactions with Combinatorics problems?*

Procedures

This section presents a brief overview of the data used in the study and the procedures employed to address each research question.

Research Question 1

To answer the first two research questions, this study uses a public dataset from a large-scale nationwide assessment in Brazil called the *Exame Nacional do Ensino Médio* (ENEM). The ENEM is a yearly exam taken by over three million students to assess their knowledge in various

subjects and is the primary criterion for undergraduate admission in Brazil. This study uses publicly available data for all years in which the exam included a Mathematics-specific section, from 2009 to 2020. In total, 28 Combinatorics problems were identified across 12 exam years. The classification process (i.e. defining whether each exam question was a Combinatorics or a non-Combinatorics problem) included the researcher's classification, a comparison with classification lists provided in ENEM preparation websites, and validation from two experienced Brazilian mathematics teachers. Mathematics problems in the ENEM exam are close-ended questions with five possible answer choices. Data on students' answer choices and responses to a sociodemographic questionnaire are publicly available on Brazil's Ministry of Education website.

The first research question inquires how the difficulty and accessibility levels of Combinatorics problems compare to problems in other domains. Difficulty is measured in this study as the rate of incorrect responses for Mathematics problems in the ENEM exam. The difference in difficulty levels between Combinatorics and non-Combinatorics problems is estimated using three regression models with increasing levels of control for potential confounding effects. The first model tests whether Combinatorics models present higher rates of difficulty than non-Combinatorics problems, without considering additional control variables. The second model includes controls for year fixed effects, and the third model adds controls for each question's position in the exam booklet.

Accessibility is defined in this study as the extent to which students understand what problems ask and engage in productive mathematical thinking while trying to solve them, without having to resort to much pre-requisite mathematical knowledge. Although this variable is not directly observable from the ENEM data, I construct a testable hypothesis that I argue to

be a corollary of this definition. If working on Combinatorics problems relies less on prior knowledge than other mathematical domains, then previous educational inputs should give students less advantage when tackling Combinatorics problems than when tackling problems in other domains. To investigate this, we can study the documented disparities in learning outcomes among students in Brazil who have differing levels of access to high-quality mathematics learning resources. A clear illustration of this difference in access to learning resources has been documented in studies comparing public and private schools in Brazil. Typically, private school students have access to superior educational resources and tend to score higher on standardized math exams. This gap has been documented at the elementary, middle, and high school levels (Cavalcanti et al., 2010; Oliveira et al., 2013; Souza et al., 2018). Based on this information, I propose the following hypothesis to investigate whether Combinatorics is a more accessible domain in Mathematics: *If Combinatorics is more accessible than other domains tested in ENEM, then the test score gap between private and public schools will be smaller for Combinatorics problems than for non-Combinatorics problems.*

Using the ENEM data, this dissertation empirically tests the hypothesis that the school achievement gap between private and public school students - a *proxy* variable for accessibility - is smaller in Combinatorics problems than in non-Combinatorics problems. I also define an alternative proxy based on the well-documented racial achievement gap that exists in Brazil. On average, white students have better access to learning resources and higher achievement rates on standardized exams than black students (Madeira & Rangel, 2014; Becker & Arends-Kuenning, 2020). The statistical tests for the proxies of accessibility are run with similar specifications as the ones used to test difficulty levels. In the case of accessibility, though, I discuss in detail the potential confounding effect of an additional control variable - the *difficulty* level previously

estimated for each mathematics problem. To control for this potential confounding effect, I resort to an additional regression specification, as well as to matching sample techniques.

Research Question 2

The second research question is addressed by investigating predictors of difficulty and accessibility of the ENEM problems using a quantitative approach. The study examines two quantifiable characteristics derived from the mathematics education literature: (i) item length, and (ii) the presence of mathematical notation. Item length is calculated using a word-counting automated software, while the presence of mathematical notation is human-coded as a binary variable (1 if there are any complex terms or symbols and 0 otherwise). The list of symbols considered complex was reviewed with the help of high school mathematics teachers.

To answer Research Question 2, these two quantified characteristics (item length and item notation) are added to the regression equations used in Research Question 1. The results of the new specification models are analyzed in two ways. First, I check whether the inclusion of the item-specific variables “explains away” eventual differences in difficulty and accessibility between Combinatorics and non-Combinatorics problems (i.e. once the additional variables are taken into account, do difficulty and accessibility levels become statistically similar across Combinatorics and non-Combinatorics problems?). Second, I analyze the sign and significance of each variable’s coefficients in the regressions (i.e. is item length a predictor of difficulty and accessibility levels in ENEM problems? Is the presence of notation a predictor?).

This study also examines the differential impact of the two quantified linguistic factors—item length and item notation—on Combinatorics versus non-Combinatorics problems. To achieve this, the empirical method previously outlined is replicated separately for Combinatorics

and non-Combinatorics problems, aiming to discern any distinct patterns these factors might present across the two categories of problems.

Research Question 3

The third research question deepens the analysis into an explanatory level, asking *how* students understand and engage with Combinatorics problems, and what prevents them from arriving at correct responses. To answer this, I conducted task-based interviews with Brazilian students, selecting problems from the ENEM exam that were part of the dataset analyzed in the previous two research questions. The study focused on students with disadvantaged school backgrounds. Interviews with 13 students were held individually via Zoom video call in September 2022, with data collected through questionnaires, video recordings, transcripts, field notes, and pictures of students' written works.

The interview design aimed to uncover students' interactions with ENEM problems, focusing on their understanding of what each problem asked and on the mathematical thinking they used in their solution attempts. Twelve problems were selected for the interviews, including eight Combinatorics and four non-Combinatorics questions. The initial themes for analysis were derived from the literature review developed in the next chapter. Following the general guidelines of grounded theory (Charmaz, 2014; Creswell & Poth, 2019), new coding themes were developed during the data analysis process to highlight emergent, unpredicted phenomena pertinent to the research questions.

Chapter II: Literature Review

This chapter presents the conceptual and empirical literature that underlies the present work. First, I introduce the mathematical domain of Combinatorics. Second, I discuss how the mathematics education literature describes Combinatorics as a domain that is simultaneously accessible and difficult. Last, I provide contextual information about education in Brazil and how Combinatorics is present in the country's most significant high-school examination - which is used as a primary data source for this study.

The Domain of Combinatorics

As mathematics has evolved over time, it has become increasingly specialized, with different domains (branches, areas) focusing on specific aspects of mathematical knowledge. The *Princeton Companion to Mathematics*, a book devoted to introducing a “large and representative sample of the ideas that mathematicians are grappling with at the beginning of the twenty-first century” (Gowers *et al.*, 2008, p.ix) defines, for example, 26 different branches of mathematics - from algebraic geometry to computational number theory, to stochastic processes. In K-12 mathematics, breakdowns of mathematics into branches are usually more straightforward. The Common Core State Standards for high school mathematics define only seven domains - number and quantity, algebra, geometry, functions, modeling, and statistics and probability.

Mathematicians and mathematics educators generally agree that this separation is somewhat arbitrary and that mathematical ideas overlap across domains. Gowers *et al.* (2008) explain that their rationale for defining mathematical branches considered the mathematical objects studied, the types of questions addressed, and the techniques used in each branch. The Common Core State Standards, on the other hand, define mathematical domains as large “groups

of related [learning] standards” (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010, p.5).

Mathematics education scholars have stressed the importance of research about teaching and learning in specific domains of mathematics (Cai et al., 2019). Algebra and Geometry are two examples of areas where research has been relatively successful in understanding domain-specific issues related to teaching and learning mathematics. In Geometry, extensive research has, for example, characterized student developmental levels of geometrical thinking (e.g., Burger & Shaughnessy; 1986, Gutiérrez et al., 1991; Van Hiele, 1986), geometric habits of mind that can support student learning in this content area (e.g., Bülbül & Güler, 2021; Driscoll et al., 2017), conditions under which the Dynamic Geometric Software can improve Geometry instruction (e.g., Hegedus & Moreno-Armella, 2009; Žakelj & Klančar, 2022), and the role of Geometry in supporting student learning of proofs (e.g., Burger & Shaughnessy, 1986; Sinclair & Robutti, 2012). In Algebra, extensive research has been done about how students interact with foundational algebraic concepts (an area known as *Early Algebra* - e.g., Blanton et al., 2015; Carraher & Schliemann, 2007, 2020), algebraic habits of mind that can support student learning in this content area (e.g., Blanton & Kaput, 2015; Driscoll, 1999), and the role of Algebra courses in defining students pathways in post-secondary education (e.g., Kim et al., 2015; Stein et al., 2011), among many other topics.

Less research has been done on how students learn and interact with mathematical ideas in the domain of Combinatorics. Combinatorics is a branch of mathematics that concerns the study of counting, arranging, and selecting objects. Combinatorial problems usually involve finding the number of possible outcomes for a given situation and rely on various counting techniques and concepts, such as permutations, combinations, and partitions. Combinatorial

reasoning and techniques are regarded as the backbones of other areas of mathematics, such as probability and computer science (Kapur, 1970; Piaget and Inhelder, 1951/1975; Tucker, 2012). The literature about different domains of mathematics suggests we might learn something about how students interact with mathematics by looking more carefully at the particularities of a domain.

In addition to its theoretical applications, Combinatorics is frequently connected to real-world problems. Many of the problems in Combinatorics found in k-12 textbooks are based on contextualized scenarios and can often be recognized by the phrase “in how many ways can you...”. One classic example is found in George Polya's famous book, "How to Solve it" (1954):

In how many ways can you change one dollar? (The “way of changing” is determined if it is known how many coins of each kind - cents, nickels, dimes, quarters, half dollars - are used) (Polya, 1954, p.238)

This problem illustrates the dual nature of Combinatorics, which is the focus of this dissertation. First, problems in this domain tend to be easy to understand. The problem statement is written in plain English, without using any challenging mathematics-specific concept or notation. Second, problems in this domain can be surprisingly challenging to solve. In the dollar problem, organizing and keeping track of all possible combinations of coins that add up to 100 cents is surprisingly demanding. When I showed this problem to a middle schooler, she delved into a compelling entry point: the simplest way to make a dollar is to use a single type of coin. We can use one hundred cent coins, or twenty nickels, or ten dimes, and so forth. Using this strategy, we find five ways to make a dollar - one for each type of coin. But we can also use two types of coins together to make a dollar - cents and nickels, for example. Counting how many ways this can be done is significantly more challenging than using a single type of coin. Focusing on the nickels can help illustrate a possible strategy for this counting. One can realize

that all nickels-and-cents ways to make a dollar can be constructed with the following procedure: pick one nickel and fill the remaining 95 cents with cents coins; pick two nickels and fill the remaining 90 cents with cents coins... until we reach the case of using 19 nickels and 5 cents. Thus, there are 19 ways to make a dollar using both (and only) cents and nickels. The final solution to the problem must include all other possibilities that use two types of coins to make a dollar, as well as all combinations that use three, four, or all five types of coins altogether.¹

Compared to other branches of school mathematics like algebra and geometry, the mathematics education literature has been relatively modest in producing empirical studies on Combinatorics. Nonetheless, previous research has explored student learning of Combinatorics at the elementary (English, 1991,1993), middle (Piaget & Inhelder, 1951/1975; Tillema, 2018; 2020), high school (Batanero et al., 1997; Lemanna et al., 2022), and post-secondary levels (Lockwood & Gibson, 2015; Lockwood & Purdy, 2019; Mashiach Eizenberg & Zaslavsky, 2004).

Combinatorics: A Paradoxically Accessible and Difficult Domain?

On the one hand, research suggests that Combinatorics is a domain that is *accessible* to students from varied backgrounds and proficiency levels in mathematics (DeBellis & Rosenstein, 2004; Goldin, 2004; Hurdle et al., 2016; Lockwood et al., 2020). On the other hand, research has also documented students struggling to solve Combinatorics problems, implying that it is a *difficult* domain to teach and learn (Batanero et al., 1997; Dubois, 1984; Hadar & Hadass, 1981). This section explores how these seemingly contrasting claims can be compatibilized. To do that, first, I define the terms *accessibility* and *difficulty* from a theoretical perspective and then briefly

¹ Alternatively, those familiarized with combinatorial methods may attempt to use a generating function to solve this problem. See Erickson (2010) for a through discussion of this problem and its generalization.

illustrate how the literature describes Combinatorics problems as both accessible and difficult. The subsequent sections in this chapter analyze the empirical results from the Combinatorics mathematics education literature in more detail and highlight their connections to issues of accessibility and difficulty.

To shed light on the apparent contradiction that Combinatorics is simultaneously accessible and difficult, I resort to the notions of *product* and *process* in students' interactions with mathematics problems. The interest in *processes* is widespread in the mathematics education literature. Indeed, some of the most influential books in the field are devoted to understanding the processes involved in solving mathematical problems (Polya, 1954; Schoenfeld, 1985). For example, Cuoco et al.'s (1996) seminal article on mathematical *habits of mind* defends a high-school curriculum focused on providing students with opportunities to “adopt some of the ways that mathematicians *think* about problems” (p.376). Their focus is not on the final answers provided by students but instead on how they get there. The processes valued by the authors include the very initial step of understanding a problem, consisting of “bringing mathematical meaning to problems and statements through definition, systematization, abstraction, or logical connections making” (p.51). Also, the National Council of Teachers of Mathematics (NCTM) played an important role in publicizing the interest of the mathematics education community in mathematical processes with the inclusion of *process standards* in their *Principles and Standards* publication (NCTM, 2000). The process standards include problem-solving, reasoning and proof, communications, connections, and representation. The five process standards that appeared in the original NCTM standards document would later influence the Common Core State Standards for Mathematics published in 2010, which includes eight

mathematical practices of similar nature (NCTM, 2000; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010).

The notion of mathematical *processes* discussed above can help understand the claims that Combinatorics problems are particularly accessible. This domain is described in the mathematics education literature as one in which students can understand what problems ask and engage in solution attempts that give rise to valuable mathematical thinking. This interpretation is expressed by Lockwood et al. (2020):

This is what we mean when we say that a counting problem is accessible – students are able to understand what a question is asking, they can begin a solution path by symbolizing outcomes, and, in the process of symbolizing these outcomes, they can engage with and discuss key combinatorial ideas. (p.5)

While *attempting* a solution, students may engage in valuable mathematical processes even if they do not find a correct answer. The literature highlights that in Combinatorics problems, students' solutions attempts frequently involve valuable mathematical processes - such as those defined by Cuoco et al.'s (1996) habits of mind or NCTM's (2000) process standards. Combinatorics problems are thus referred to as accessible in the sense that they “lend themselves to desirable practices such as generalizing [...], proving, using multiple representations, and attending to precision” (Sandefur et al., 2022, p.756).

In contrast, a product perspective focuses on students' final answers. For example, studies that explore inequities in student achievement based on large-scale exams can be seen as mostly product-oriented, as they are concerned with how many (and which) students answer problems correctly. Examples of this literature include investigations of racial and gender gaps in student achievement (e.g., Darling-Hammond, 2014; Lubienski, 2008), and the role of socioeconomic variables as mediating factors or predictors of mathematics achievement (e.g., Armor et al., 2018; McGraw et al., 2006).

Claims that Combinatorics is a particularly *difficult* domain of Mathematics are usually associated with the idea that few students answer Combinatorics problems correctly. Lockwood and Gibson (2016), for example, explain that “although counting problems are easy to state and provide rich, accessible problem-solving situations, there is much evidence that students struggle with solving counting problems correctly” (p. 247). In support of this claim, research on Combinatorics has documented incorrect response rates of over 50%, even after students received direct instruction on the subject (Batanero et al., 1997). Arguably, no general criterion exists to determine what percentage of incorrect responses makes a domain “difficult”. Even so, this numerical evidence calls for attention when combined with researchers’ and practitioners’ commentaries about Combinatorics. Problems in this domain are described in the literature, for example, as “known to be difficult for students” (Lockwood and Purdy, 2009, p.225) and as “extremely challenging to solve” (Hurdle et al., 2016, p.216).

Combinatorics is thus a domain characterized in the literature as providing broad access to important mathematical thinking (process), even though students struggle to answer problems correctly (product). To simplify the vocabulary, in the remainder of this dissertation, I reserve the term *accessibility* to the process aspect of students’ interactions with mathematics problems. I use the term *difficulty* to refer to product aspects of students’ interactions with problems. To illustrate this distinction, Figure 2.1 describes problems with different levels of difficulty and accessibility.

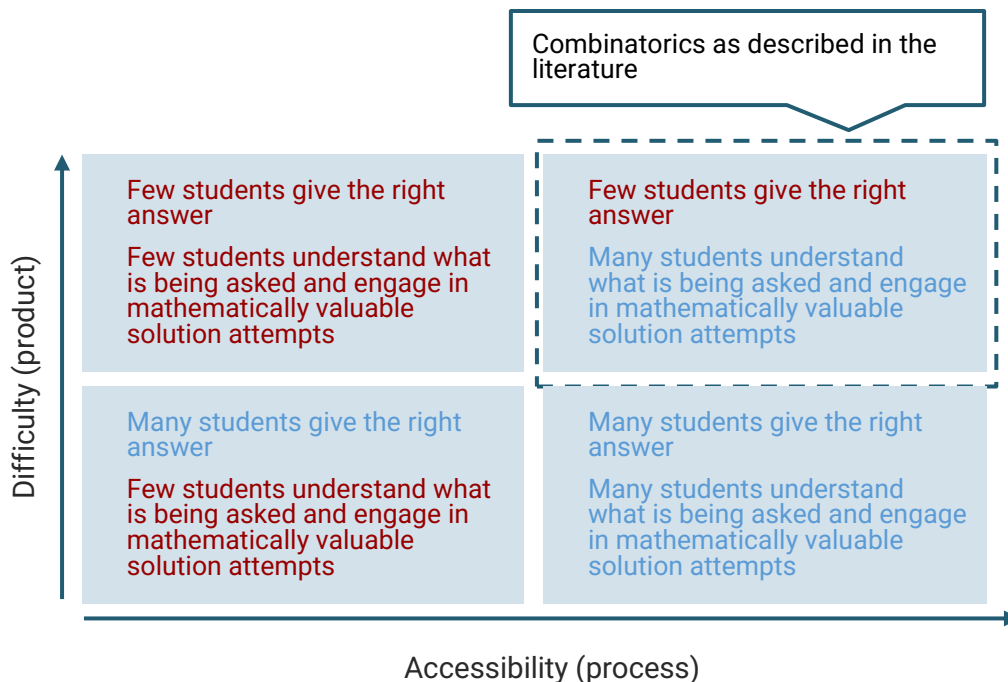


Figure 2.1: Difficulty and Accessibility in Mathematics Problems

By investigating issues of accessibility and difficulty, this study is ultimately concerned with whether and how the domain of Combinatorics can contribute to more equitable learning in mathematics. Although helpful in organizing a theoretical framework in this dissertation, the definitions of accessibility and difficulty are not without limitations regarding the multiple aspects that equity can entail. Gutiérrez’s (2007, 2008) framework of equity is instrumental in highlighting what the above definitions encompass and what they leave out. In Gutiérrez’s (2007, 2008) framework, there are four dimensions of equity: access, achievement, identity, and power. The analysis proposed in this dissertation relates only to the first two of those four dimensions.

In Gutiérrez’s framework, *access* focuses on students’ opportunities to learn mathematics, thus on the “resources available [for students] to engage with quality mathematics” (Gutiérrez, 2008, p.360). While her definition focuses on the availability of resources that enable students to engage with mathematics in meaningful ways, it closely relates to the definition of

accessibility proposed in this study (the extent to which students can understand what a problem is asking and engage in valuable mathematical processes while attempting a solution), which is concerned with the *de facto* engagement of students with valuable mathematics processes. Gutiérrez's (2008) definition of the *achievement* dimension of equity concerns "tangible results for students at all levels of mathematics" (p.3). The achievement dimension in her definition includes but is not limited to standardized scores - it encompasses, for example, student participation in class and which mathematics courses students take throughout their lives. Her *achievement* dimension thus encompasses this dissertation's definition of difficulty (the extent to which students answer mathematics problems correctly). The third and fourth dimensions of equity stressed by Gutiérrez, *identity* and *power*, are not directly addressed in the present work. While *identity* regards students' sense of themselves as capable and successful mathematics learners, *power* relates to the role that mathematics plays in society and the need to challenge existing power structures that perpetuate inequities.

Combinatorics as an Accessible Domain

This subsection reviews empirical studies that documented learners engaging in valuable mathematical thinking while working on Combinatorics problems. Combined results from these studies provide evidence that Combinatorics problems can be accessible to students at different age levels. To illustrate these results, this section highlights episodes that showcase students understanding what Combinatorics problems ask and engaging with them in mathematically valuable ways.

A landmark study in the learning of Combinatorics is the work of Piaget and Inhelder, who interviewed elementary school students engaging with combinatorial tasks (Piaget & Inhelder, 1951/1975). Their research is centered on interpreting the development of children's

ability to perform combinatorial thinking, as part of their extensive effort to document children's developmental stages. The tasks included in the study asked students to "permute or combine in all sorts of ways small sets of counters" (Piaget and Inhelder, 1951/1975, p.162). Their report includes student quotes that help illustrate how young children interact with combinatorics problems. In the following passage, Lau, a 12-year-old student, is asked to find how many different pairs could be formed with four colors of counters (coded in the report by the letters A, B, C, and D)²:

'Can you find a way of doing all of them, forgetting none?' [interviewer.] 'Yes' (He puts together AB, CD, AC, BD, AD, and BC.) 'Are you just guessing?' 'Not at all. I looked for four [counters] of each [color] and saw that three were enough. There would be four if we put yellow with yellow, blue with blue, etc' (Piaget and Inhelder, 1951/1975, p. 171-172. Italics in the original).

In this passage, the student shows evidence of having anticipated the structure of the final solution before engaging in its construction. This mathematical process is highlighted in the seventh Common Core Standard for Mathematical Practice - "Look for and make use of structure" (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). In particular, he seems to invoke the notion that the set of all possible outcomes must be symmetric across the colors (each color would need to appear in a pair with each of the remaining three colors). Looking for symmetries is precisely one of the problem-solving heuristics emphasized in Polya's influential book (1954). The student also spontaneously delves into an extension of the problem, anticipating what would happen if he were also to include pairs of counters of the same colors - engaging in a process that resembles the

² This problem involves *combinations* - structures in which the order of the elements is irrelevant. The child's answer implicitly indicates that he assumed that pairs of counters of the same color were not allowed as solutions. Taking that restriction into account, the set of outcomes in this problem is {{A,B}, {A,C}, {A,D}, {B,C}, {B,D}, {C,D}}. There are thus six different pairs of counters to be created.

mathematical habit of mind of *tinkering*, wondering “what happens if something is left out or if the pieces are put back in a different way” (Cuocco et al., p. 379).

Another episode of student reasoning focused on the structure of the problem appears in the second combinatorial task discussed in the book, which deals with arrangements. In this case, the authors explicitly told students that the order in which the objects were arranged mattered. They provided students with counters to form different ordered pairs (or triads or quadruples) by “placing the necessary colors in a row” (Piaget and Inhelder, 1951/1975, p.175). In the following passage, Lay, a 12-year-old student, explains how many different rows can be made with counters of three colors (no color repetition allowed)³:

Lay (12; 4) is moving around three counters A, B, C (with no other counters present so as to avoid empirical groupings). “Can we predict how many different ways we can arrange them?” [interviewer]. “*With three counters there ought to be six different ways because [with color A in the first position] each one [of the remaining counters] can change places twice.*” (Piaget and Inhelder, 1951/1975, p.189-190. Italics in the original).

Similarly to the previous case, the student in this passage was able to reason about the structure of the problem. More specifically, they unveiled the multiplicative structure of the set of possible outcomes (known as the “multiplication principle” in Combinatorics) and provided a compelling explanation for their answer. When asked to solve the problem with four colors available, the student first guessed that the answer would be *eight*, which is incorrect. But after being challenged by the interviewer that eight might be too few possibilities, the student focuses on the multiplicative structure of the problem once again: “*It’s simple. I started with the same*

³ Similarly to the previous passage, the child’s answer implicitly indicates that they assumed that repeated counters of the same color were not allowed in solutions. Taking that restriction into account, the set of outcomes in this problem is {ABC, ACB, BAC, BCA, CAB, CBA}. There are thus six different ordered triads of counters to be created.

color (A). That made six. Therefore, twenty-four in all because for each color, we get six, and six times four is twenty-four.” (Piaget and Inhelder, 1951/1975, p.190. Italics in the original.)

Aligning with the Piagetian tradition, English (1991) also interviewed students to document children’s developmental stages of combinatorial reasoning. As she explains, her task was intentionally more contextualized than Piaget and Inhelder’s (1951/1975) original problems. The inclusion of a contextual background was intended to make the task more accessible to young children. English’s (1991) task asked students how many different outfits they could create for a bear. For this, the students had available manipulatives representing tops, pants, and skirts in varying quantities. In the most complex version, students had three types of tops and three types of skirts available. English’s (1991) intentional design seems to have indeed allowed younger children to provide mathematically valid justifications for their solutions to the proposed tasks when compared to Piaget and Inhelder’s (1951/1975) study. In English’s (1991) interviews, children aged 5 to 9 years were able to express a type of combinatorial solution strategy that she named “odometer”. An odometer strategy is one in which one object is held constant (e.g., tops) and gets successfully combined with each available option of the other clothing piece (e.g., pants). This reasoning was expressed by students when asked how they could make sure they were finished creating all possible outfits, as exemplified in the following passage:

‘I can't make any more outfits because I've matched each top with each pair of pants, and there are no more tops’ or, ‘I know I can't make any more outfits because there are three different tops and I've used each pair of pants three times.’ (English, 1991, p.461)

Similar evidence of productive mathematical engagement in Combinatorics is provided by Maher et al. (2010) in a longitudinal study that followed a cohort of students for over 12 years, from first grade to college. The study’s main interest is not in Combinatorics itself but instead in the development of students’ autonomy in reasoning and argumentation. The

researchers decided, though, that Combinatorics would be a promising domain for this goal and organized a series of Combinatorics problems that were presented to the students over the years. Their choice to work on this domain is explained, at least in part, by their judgment that Combinatorics would be an accessible domain. According to the authors, this domain of mathematics was chosen because “in working on these problems, students can find the need to organize their work systematically, look for patterns, and generalize their findings” (Maher et al., 2010, p.9). Most informative is the authors’ commentary on the progression of student mathematical thinking during the study’s years:

In the early years of the study, children began to use inductive reasoning, to organize work by cases, and to think about justification through contradiction. By middle school, these forms of reasoning were more sharply defined, and other forms of reasoning emerged, such as controlling for variables. In high school, students began the process of building isomorphisms, using their own notation as well as standard notation to describe how some problems were related to each other and ultimately to Pascal’s triangle.” (Maher et al., 2010. p. 24).

An illustration of such mathematical development provided in their study involves a second grader working on a dressing problem similar to the task used by English (1991). There were three types of t-shirts and two types of jeans available, and students were asked to find how many different outfits could be created. In the following passage, a student explains her thinking about how to check whether her final answer was correct. As in one of the passages from Piaget and Inhelder (1951/1975) discussed above, her reasoning uses the multiplication principle:

“Dana, early on, expressed verbally her understanding of the structure of the solution when she indicated that each of the three shirts could be combined with each pair of jeans. She said, “He can make all three of these shirts with that outfit” (Martino, 1992, p. 47). It can be concluded from this statement and the

subsequent problem-solving steps that she took that she had built a scheme that closely matched the problem solution.” (Maher et al., 2010. p. 21).

In conclusion, the literature has documented that Combinatorics can be an accessible domain for learners of different age groups. Through empirical studies, researchers have observed students engaging with valuable mathematical thinking while working on Combinatorics problems. The accessibility of this domain is also supported using contextualized tasks, allowing even younger students to interact with fundamental combinatorial concepts such as the multiplication principle. As students advance in their mathematical education, their engagement with Combinatorics problems seems to help them refine their reasoning and argumentation skills.

Linguistic Factors and Accessibility

One explanation provided in the literature for why Combinatorics problems are accessible is that they do not require much mathematics-specific vocabulary or symbols to be stated. This is frequently stressed by authors commenting on the nature of discrete mathematics - the broader area of mathematics that encompasses combinatorics. Discrete mathematics is described, for example, as an area in which “the problems themselves do not require knowledge of technical definitions or specific mathematical knowledge, and students can exemplify and explore objects” (Sandefur et al., 2022, p.756). Lockwood et al. (2020) illustrate this characteristic of Combinatorics with an example of a typical problem in this domain:

You have twenty different books, and you want to take three of them with you on vacation. How many different possibilities are there for which three books you could bring? Note that the vocabulary is easily understood – there are no technical mathematical terms, which would require previous definition (e.g., “codomain”) (p.4)

Indeed, many of the problems used in the empirical literature on the learning of Combinatorics are stated using only “common language”, without any specific mathematical symbol. This is the case, for example, of the counters tasks (Piaget and Inhelder, 1951/1975) and bear-dressing tasks (English, 1991) discussed above. Even studies that explored the mathematical thinking of college students in Combinatorics have used lists of problems composed only of words and natural numbers. For example, Lockwood and Gibson (2016) used ten problems to explore undergraduates’ use of lists in solving combinatorics problems. Apart from English language words, they only include natural numbers, which are frequently written as words, as in the following example:

There are five different Spanish books, six different French books, and eight different Russian books. How many ways are there to pick a pair of books that are not both in the same language? (Lockwood and Gibson, 2016, p.253)

For the sake of preciseness, one can point out that two problems in their list use letters that could, in principle, be interpreted as complex mathematical symbols. In those two instances, though, the letters provide only the context for “what is to be counted”. Thus, in this case, the letters can be interpreted by students as any other discrete object or characteristic, such as a t-shirt or a color. Arguably, then, students can understand what the problem asks without the need for complex prerequisite mathematical knowledge. This contrasts with other uses of letters in mathematics, such as to represent an unknown quantity or a relationship between two quantities in algebra problems. The “concrete” use of letters in Combinatorics is exemplified in the following problem:

“How many arrangements are there of the letters A, B, C, Z, Z, Z, Z, where the A, B, and C occur alphabetically (they do not have to appear together as a group)?” (Lockwood and Gibson, 2016, p.253)

Studies outside the literature on Combinatorics provide additional evidence that linguistic factors can play an essential role in how students interact with mathematics problems. For example, Avalos, Medina, and Secada (2018) interviewed 8th graders solving mathematical problems. The authors found that students who were less proficient looked for mathematics-specific symbols and vocabulary “strictly for procedural clues on how to solve the problems” (p. 70). Conversely, their high-achieving peers used Mathematics-specific symbols and vocabulary as references for sense-making. Thus, the presence of technical symbols seems to influence access by leading some students - particularly those with lower proficiency levels - away from understanding the mathematical meaning of what is being asked in a problem. Consequently, a lower presence of mathematics-specific symbols and vocabulary may be one of the factors that allow Combinatorics problems to be accessible to students of varied proficiency levels.

Variety of problem-solving approaches and accessibility

Another explanation for why Combinatorics problems are accessible is that students can approach them from multiple perspectives. While the linguistic factors discussed previously influence how students with varied backgrounds understand what problems in Combinatorics ask, this additional argument relates to students’ mathematical thinking *after* they have understood what the problem is asking.

As an example, consider the simple case of the bear-dressing problem proposed by English (1991), with three types of pants and three types of tops. Students can start by holding a type of pants fixed and figuring that it can be combined with each of the four tops to create a different outfit. The student can either repeat the process to realize there are twelve outfits ($4+4+4=12$), or use the multiplication principle - as young students in the highlighted passages have -, to arrive at the same result ($3 \times 4 = 12$). Those two alternative routes can also be reproduced

by starting with a fixed top instead of fixed pants (e.g., $3+3+3+3=12$; $4x3=12$). If sufficient manipulatives are available, a student can also solve the problem by physically recreating all possible outfits and then backward-testing that all cases are already included (for example, by checking whether each pant has already been combined with all available tops, and vice-versa).

Problems in Combinatorics are thus regarded as accessible because once students understand what the problem is asking, they can use varied background knowledge to engage in solution attempts (English, 2005). On this matter, English (2005) explains that “combinatorial problems lend themselves to a variety of solution approaches, enabling children with minimal content knowledge to work towards a solution” (p.133).

Further, the literature stresses that by allowing for multiple solution strategies, Combinatorics problems also provide fertile grounds for students to develop the ability to communicate their mathematical reasoning. When students create varied personal strategies to solve problems, they can explain, compare, and connect their approaches. Combinatorics problems can thus provide “a context for discussions about the *why* and *how* of a mathematics problem, as well as allowing for students to support their ideas.” (Hurdle et al., 2016, p.221). This perspective aligns with the broader literature that discusses the characteristics of mathematics tasks that support students in engaging in rich mathematical discussions.

Schoenfeld (2020), for example, explains that:

Problems with multiple entry points allow more students to find handholds into the mathematics and are the seeds for rich student-to-student discourse, which is essential (cf. “group-worthy problems,” Cohen & Lotan 1997) and seeing the connections between multiple approaches deepens one’s understanding. (p.1170)

In sum, when students solve a problem using different personal approaches, they have rich mathematical material to talk about with each other. Combinatorics is arguably a fruitful domain to generate such varied solution paths.

Combinatorics as a Difficult Domain

Research also highlights that Combinatorics problems are generally difficult to solve. To clarify the distinction between difficulty and accessibility, in this study, I defined difficulty simply as the rate of incorrect response provided in a problem. In direct relationship with this definition, several studies have documented high rates of incorrect responses among students working on combinatorics problems. Also related to the issue of difficulty, studies have delved into the different kinds of mistakes incurred by students in trying to solve Combinatorics problems.

Batanero et al. (1997) investigated difficulty levels across different types of Combinatorics problems. Their study took place in Spain and involved 720 students aged 14 and 15 who answered a test with thirteen introductory Combinatorics problems. Part of the students received instruction in Combinatorics before answering the problem set, and part did not. Among the sample of students who had received specific instruction in Combinatorics, the authors encountered a 59.1% incorrect response rate across problems (78.3% among students with no instruction on the subject matter). The main conclusion of the study was that the “implicit combinatorial model” embedded in the problems (e.g., selection, distribution, partition) was significantly associated with the rates of incorrect responses. Among the problems in their questionnaire, partition problems had the highest rates of incorrect responses (75.8% across all students). The authors also classified fourteen different types of errors found in student reasoning. The most common source error was the confusion of two important concepts in combinatorics: combinations and arrangements. This mistake would occur when students fail to “distinguishing the order of the elements when it is irrelevant or, on the contrary, not considering the order when it is essential” (p.192).

Lamanna et al. (2022) used a similar approach to Batanero et al. (1997) to study the effects of instruction on Italian students' combinatorial strategies. Their sample comprised 115 students aged 15 to 18 years old who worked on a test with three Combinatorics problems adapted from Navarro-Pelayo (1994). Although the problems varied in their underlying combinatorial concepts (distribution, selection, and partition), they shared one characteristic - all three problems had numerical answers below 10. As the authors explain, they intentionally chose problems with small numerical answers, which allowed students to "produce a complete list of configurations" (p.7). The authors report a 42.7% incorrect response rate to Combinatorics problems among students who received instruction in Combinatorics (65.6% across students with no instruction). In their analysis of students' mistakes, the authors reveal that even though the problem could be solved by listing all possible cases, students made frequent and varied mistakes while attempting to use this strategy. In the distribution and partition problems, many students created lists that were conceptually wrong, which included objects that should not be counted. These students overestimated the problems' final answers. In selection problems, a different type of mistake was common among students who tried to use lists. In this case, students frequently failed to account for all cases, even when their lists were organized in ways that indicated a systematic organization of the counting process.

In addition to empirical studies of students' errors, the literature characterizing Combinatorics as a difficult domain includes essays and commentaries in which the authors build on their teaching experiences. Annin and Lai (2010), for example, use an illustrative problem involving permutations and combinations to describe three different types of mistakes incurred by students. The problems they use - along with nuanced variations discussed in the paper, are targeted at high school-level Combinatorics. They introduce the first problem with the

question: “How many flags of 8 horizontal stripes contain at least 6 blue stripes if each stripe is colored red, green, or blue?” (Annin & Lai, 2010, p. 404). Based on this problem, the authors discuss an incorrect use of the multiplication principle. In their example, a straightforward (albeit incorrect) application of the multiplication principle leads to the double counting of some flags. The two other mistakes discussed in the study also involve some form of conceptual confusion as to how and when to take the order of the counted objects into account. In the second example, students mistakenly impose order in a problem in which order should not be taken into account. In the last discussed type of mistake, the converse happens. Students fail to consider order in a problem in which order is relevant, leading them to understate the final answer. A takeaway from Annin and Lai’s (2010) discussion is that finding correct answers in Combinatorics can depend strongly on navigating nuances of problems.

In another essay that describes possible student mistakes in Combinatorics, Hadar and Hadass (1981) discuss a problem known as the “Bernoulli-Euler Problem of the Misaddressed Letters” (p. 435). This problem is significantly more challenging than the problems used by Annin and Lai (2010) and is present in post-secondary level Combinatorics textbooks (e.g., Andreescu and Feng, 2003; Cameron, 1994). The problem is stated as follows:

“Someone writes n letters and writes the corresponding addresses on n envelopes. How many different ways are there of placing all the letters in the wrong envelopes?” (Hadar & Hadass, 1981, p.435)

Based on this problem, the authors describe seven examples of possible student mistakes. Pitfalls ranged from conceptual challenges, such as misinterpreting what is being counted, to calculation mistakes - such as failing to count all cases. Albeit at a more complex level, these mistakes mirror the pitfalls described by Annin and Lai (2010) for secondary school-level

Combinatorics. Again, they highlight the importance of navigating the details to correctly answer problems in this domain.

The difficulties faced by students in distinguishing between different combinatorics problems were further investigated in higher education settings. Lockwood et al. (2018) were particularly interested in whether undergraduate students use a specific combinatorial method (binomial coefficients) to solve two different kinds of problems, which they refer to as Type I and Type II problems. Type I problems are problems in which the context directly involves the selection of k among n distinguishable objects, with the order of selection being irrelevant. Type II problems are problems in which the order matters, but not all objects being ordered are distinguishable. While both types of problems can be solved using binomial coefficients, their results showed that even students with prior experience with combinatorics were significantly less likely to use this combinatorial method to solve Type II problems correctly. Wasserman and Galarza (2019) also investigate how post-secondary students (in this case graduate students) navigate nuances between two Type I and Type II problems and provide further insight into where these difficulties may come from. Analyzing students' attempts to explicitly connect both types of problems to the binomial coefficients, the authors provide evidence that students view Type I and Type II problems differently. The authors highlight that students tend to associate Type II with a set of outcomes that is not directly connected to the concept of *combinations*, even though a different and also valid set of outcomes is. Together, these studies suggest that the way students think about the sets of outcomes of combinatorics problems can be a key factor in determining how they navigate problems' nuances.⁴

⁴ For a suggestion on how current combinatorial notation can be expanded allow for more explicit treatment of sets of outcomes in combinatorics instruction, see Wasserman (2009).

Linguistic factors and difficulty

As discussed in previous sections, the Combinatorics education literature has emphasized the non-technical language of problem statements as an explanatory factor for accessibility. Less attention has been devoted to the possible connections between linguistic characteristics and difficulty levels in Combinatorics problems. Thus, I turn to the more general mathematics education literature - outside the realm of Combinatorics - to identify some possible connections between linguistic characteristics of mathematics problems and their difficulty levels.

In their systematic review of linguistic factors contributing to difficulty in elementary school word problems, Daroczy et al. (2015) differentiate structural factors and semantic factors. Structural factors are related to “complexity at the micro-level of word and sentence forms as well as at the macro-level of the discourse structure of the WP [word problem] passage.” (Daroczy et al. 2015, p. 4). The authors point out that, in general, higher complexity in structural factors is associated with higher problem difficulty. Examples of these constructs include the “number of letters, words, sentences, mean word, and sentence length, or the proportion of complex (long) words” (Daroczy et al. 2015, p. 4). Problems with longer statements and with more complex words and sentences tend to be more difficult for students.

Semantic factors differ from structural factors in that they take into account the meaning of words and sentences. The most important example of semantic construct discussed by the authors is the presence of explicit verbal cues that hint to students about what mathematical operation they can use to solve the problem. For example, verbs like “joining” or “gathering” tend to be interpreted by students as hints that the problem can be solved with the sum operation. As Daroczy et al. (2015) summarize in their review, the literature indicates that the absence of such verbs is associated with lower correct response rates in word problems.

Although the studies reviewed by the authors are focused on elementary school-level problems involving arithmetic operations, the results highlighting the importance of verbal cues can bring insights into Combinatorics. As Annin and Lai (2010) explain, Combinatorics problems are not easy to categorize:

“What we often find challenging about teaching students to count, however, is that most problems do not fall cleanly into one and only one standard category of counting problems.” (p.404)

By not being easily categorizable, Combinatorics word problems may be less prone to the inclusion of verbal cues. As in the case of elementary school arithmetic word problems, the absence of verbal cues might be a factor behind the high difficulty levels of Combinatorics problems.

Absence of powerful methods and difficulty

Another explanation described in the Combinatorics literature for why problems in this domain can be difficult is related to challenges in categorizing combinatorial problems and choosing appropriate methods to address them. This argument was made by Kapur (1970) in his seminal defense that discrete mathematics - including Combinatorics - should be part of the school curriculum. Compared to Calculus, for example, where some techniques can be used to solve a broad range of problems, “in the mathematics of the discrete not many such powerful methods are available and ingenuity is always required” (Kapur, 1970, p.114).⁵

Consequently, simply memorizing formulas or predetermined procedures is insufficient for solving Combinatorics problems correctly. This poses a challenge for students accustomed to tackling problems by merely applying formulas. Annin and Lai (2010) echo this interpretation:

⁵ An example of a powerful method in Calculus is L'Hôpital's rule. This technique can be used to compute limits of expressions that can be written as a division in which the dividend and divisor both tend to zero or both tend to infinity.

“Many times, no rigid procedures or formulas can be used to solve the problems directly, and students simply do not know where or how to approach them” (p.403)

This observation aligns with previous empirical findings. Batanero et al. (1997) identified the most common error in Combinatorics problems as the conceptual confusion between permutation and arrangement situations, a point reiterated by Annin and Lai (2010). Students committing this error attempt to apply a specific method - such as the formula for combinations or arrangements - to an inappropriate context.

In the remainder of this dissertation, I use the expression “absence of powerful methods” to refer to students’ difficulties in identifying combinatorial tools to use for each problem. As a clarification, I do not actually imply that no powerful methods exist in the combinatorial domain. In fact, it may be argued that mathematicians have developed many powerful combinatorial tools of varying degrees of complexity – from the multiplication principle to generating functions. The phenomenon I refer to as the “absence of powerful methods” refers thus to challenges of differentiating between types of Combinatorics problems solvable by particular methods and combinatorial tools.

Building on results and ideas from the literature on Combinatorics mathematics education, Figure 2.2 summarizes hypotheses about how some characteristics of problems in this domain are related to accessibility and difficulty.

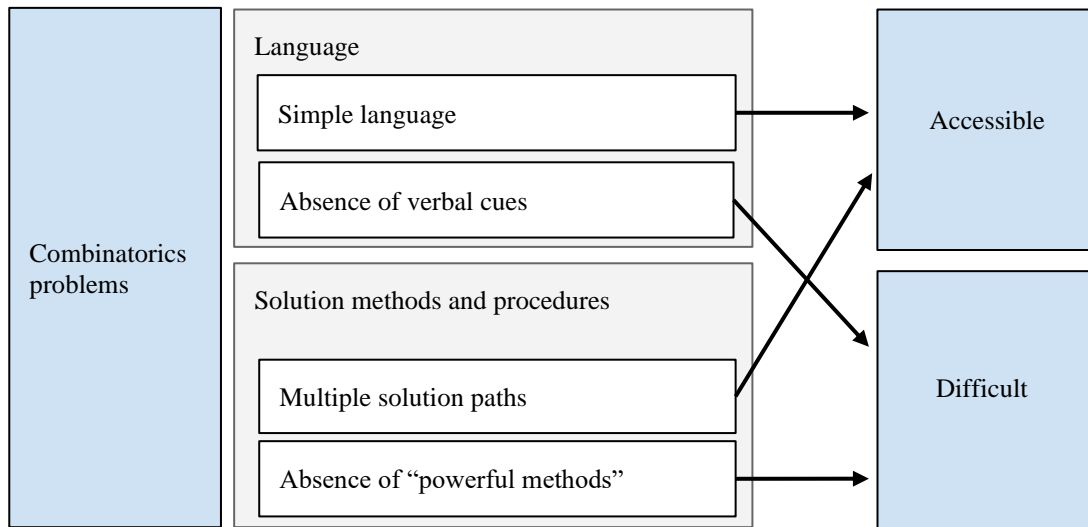


Figure 2.2. Schematic representation of hypotheses drawn from the literature.

Educational Inequalities in Brazil

Since the empirical analysis of this dissertation uses data from Brazil, I provide in this section a brief introduction to the educational inequalities and the presence of Combinatorics in the Brazilian K-12 curriculum.

Brazil is known for its extreme economic inequality. As of 2021, the average income of the bottom 50% of the Brazilian population was 29 times less than that of the top 10% - as a comparison, this figure was 17 for the United States and 7 for France (Chancel et al., 2022). Inequalities in education are less known but are equally severe. These disparities are particularly evident when examining the differences between public and private schools, as well as racial inequalities in educational achievement.

About one in five students in Brazil studies in private schools. Private schools generally have better resources and receive students from wealthier families. In general, the literature has documented higher average student achievement across private schools, in comparison to public

schools. In PISA 2018, Brazil had the largest private-public school gap among the 40 regions assessed in reading - a gap of 102 points, or just over a standard deviation (OECD, 2020). Using national 8th-grade mathematics standardized tests, Oliveira et al. (2013) estimate that public school students score on average one standard deviation below private school students. Among the variables with the highest explanatory power of scores, the authors point to teacher experience, teacher credentials, and school resources. In high school, Cavalcanti et al. (2010) estimate that standardized test scores of public high school students are on average about 17% lower than that of private school students across all subject areas. Even after controlling for a large array of individual, family, and school characteristics in multivariate regressions, the private-public does not vanish and is estimated to remain to be about 4%. In the face of this result, the authors conclude that the lower average achievement of public school students is not only attributable to a “ ‘weaker background’ when compared to private school students but also because of the low quality of public schools” (p.407).

In addition to the private-public school disparities, the literature has also documented persistent gaps between the average achievement of white and black students in standardized test scores. Using administrative data from the state of São Paulo, Madeira and Rangel (2014) provide evidence that achievement gaps in mathematics are persistent across grades. From 1st grade to 12th grade, the authors estimate achievement gaps between white and black students to range from between 17% and 30% of a standard deviation. Analyzing mathematics 9th-grade standardized scores, Becker and Arends-Kuenning (2020) find that only part of the racial gap is explained away by school resources (e.g., level of experience of teachers and principals, school location) and socioeconomic variables (e.g., parental level of schooling, whether the student works). Especially at the upper level of the grade distribution (students with higher scores), the

authors conclude that racial gaps are also attributable to “non-observable issues, such as discrimination and stereotype threats” (Becker & Arends-Kuenning, 2020, p.18).

In summary, the Brazilian K-12 education system struggles with various forms of inequities, including disparities between public and private schools and racial disparities in student learning. These inequities contribute to an uneven playing field for students, with those from marginalized groups often facing greater challenges in accessing quality education and achieving academic success. Working towards increasing access to mathematics learning is an urgent matter for the country.

Combinatorics in the *Exame Nacional do Ensino Médio*

Brazilian higher education consists of universities, university centers, and colleges, akin to the U.S. system. Public and private higher education institutions vary in quality, but differently from K-12 schools, in this case, public institutions are better on average. Public universities in Brazil are tuition-free and highly competitive, with entrance largely determined by a national entrance examination - The Exame Nacional do Ensino Médio (ENEM). Private universities, on the other hand, charge tuition fees and have less rigorous admission criteria in general.

The ENEM was introduced in Brazil in 1998 as an initiative by the Ministry of Education (MEC) to assess the quality of secondary education. Since its inception, the ENEM has evolved to become the largest standardized exam in Brazil, serving multiple purposes, including admissions into public universities and as a means for obtaining public scholarships and financing for enrollment in private higher education (Mello, 2022). The ENEM exam consists of two days of testing, encompassing a total of 180 multiple-choice questions and one essay question. The test is divided into four main areas of knowledge: Natural Sciences (Biology,

Physics, and Chemistry), Mathematics, Human Sciences (History, Geography, Philosophy, and Sociology), and Language (Portuguese, Literature, Foreign Language). In addition to the multiple-choice questions, the ENEM includes a written essay that assesses students' ability to construct a coherent argument, demonstrate critical thinking, and employ proper grammar and syntax.

As will be discussed in the next chapters, the ENEM exam has included at least one Combinatorics problem in its mathematics section every year since 2009. Since the publication of ENEM content guidelines in that year (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira [INEP], 2009) until now, the exam has maintained stable content guidelines and exam format. The ENEM exam content guidelines make two references to Combinatorics. First, under the “Competence area 1: Creating meaning for the natural numbers, integers, rationals and real numbers”, Combinatorics is mentioned in the second *habilidade* (skill) “Identify number patterns our counting principles” (INEP, 2009, p.1. Translated by the author).⁶ Second, “counting” is mentioned as one of the nine content topics under the broad area of *numbers* in the document’s Appendix section:

Numbers: operations with number sets (natural, integer, rational, and real numbers), inequalities, divisibility, factorization, ratios and proportions, percentage and interest, dependence relationships between quantities, sequences and progressions, *counting principles*. [emphasis added]. (INEP, 2009, p.9.)

The Brazilian Ministry of Education provides public access to the database of ENEM test results, including student answers to all multiple-choice questions as well as their responses to a socio-demographic questionnaire, making it a fruitful context to study Combinatorics.

⁶ All documents originally in portuguese such as curriculum guidelines from INEP and problems from the ENEM exam presented in this dissertation were translated by the author. To make the text easier to read, I omit the reference to translation authorship in the remainder of this dissertation.

Chapter III: Methodology

This chapter reports on the methods and procedures used in this dissertation. I first describe the quantitative part of the study, in which I explore data from a large-scale examination in Brazil, and then the qualitative section, which is based on student interviews. The research questions I propose in this study are the following:

- (1) How do Combinatorics problems compare to problems in other domains in terms of difficulty and accessibility?*
- (2) What (if any) are some characteristics of Combinatorics problems that are associated with their difficulty and accessibility levels?*
- (3) What are some of the ways in which students access problems in Combinatorics? What are some of the ways in which students fail to provide correct answers to problems in Combinatorics? What is the relationship between difficulty and accessibility in students' interactions with Combinatorics problems?*

Quantitative Data and Research Questions 1 and 2

Combinatorics has been part of Brazil's secondary national curriculum guidelines for over two decades (Silva, 2016). To answer the first two research questions, I use a quantitative analysis that takes advantage of a public dataset of a large-scale nationwide assessment in Brazil that includes Combinatorics problems, ENEM. The ENEM exam is taken yearly by over three million students in Brazil to assess their knowledge in Mathematics, Natural Sciences, Social Sciences, Portuguese Language, and Foreign Language (English or Spanish) at the end of high school. Most higher education institutions use it as the primary selection criterion for undergraduate admissions, making it a high-stakes test for students interested in obtaining a higher education degree. Brazil's Ministry of Education provides public access to the de-

identified microdata on all students who took the test. Publicly available data includes students' responses to the exam questions and a socio-demographic survey since the exam was created in 2004.

In this study, though, I only use data from the exams held between 2009 and 2020. In these years, the Mathematics section of ENEM had a standard structure consisting of 45 multiple-choice items. I only use information from 2009 onward because, before that, the exam had a different format in which all items were “multidisciplinary”, thus with no separation of a specific Mathematics section. For years 2016 and 2020, I also used ENEM's second application, offered only to students impeded from taking the first application due to external reasons - such as the test site shutting down due to electricity supply problems on the exam date. In these two years, the number of students taking the second application was unusually large (I defined a threshold of $n > 1000$ for including the second application in any given year).

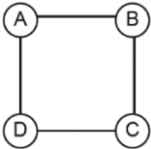
I identified which ENEM examination questions were in the Combinatorics domain with a three-step procedure. First, I read all exam questions in the selected years and classified them as either Combinatorics or non-Combinatorics items. Second, I compared my list of Combinatorics items with lists of items provided by websites dedicated to ENEM preparation.⁷ My own list was more comprehensive and included all the Combinatorics items I found on the websites' lists. Third, I showed my list of Combinatorics items to two experienced Brazilian secondary mathematics teachers and asked them to judge whether they would classify each of those problems as belonging to this domain. There was complete agreement that all problems in the list were problems in Combinatorics. With this procedure, I identified 28 combinatorics problems across the 12 exam years from 2009 to 2020.

⁷ I used the following two websites, which ranked the highest on a major search engine: qconcurso.com and m-mais.com. Access on March 1st, 2022.

While some of the problems in ENEM contained many indicators supporting their classification as Combinatorics problems, other problems were not as well defined. As a general rule, I opted to be parsimonious by classifying as Combinatoric problems only the problems that did not significantly involve mathematical concepts from other domains. The example provided in Figure 3.1 illustrates what I considered to be a “classic” Combinatorics problem. First, the problem asks for “how many different ways” something can be done, thus following a prevalent problem structure in Combinatorics. Second, the core mathematics needed to solve the problem involves working with a counting process. That is, the crucial, likely most challenging aspect of the problem involves working with combinatorial ideas. While it can be argued that solving the problem requires notions of geometry (i.e. the student needs to know the meaning of two circles being “connected by a segment”), I assumed that students could solve the problems using the literal meaning of the words “connected” and “segment” by relying on the problem’s image. I thus interpreted that the Geometry domain played a less important role in this problem.

ENEM 2016 - Question 172, blue version of the exam

To stimulate his daughter's reasoning, a father made the following drawing and handed it to the child along with three pencils of different colors. He wants the child to paint only the circles and asks that those connected by a segment must have different colors.



In how many different ways can the child do what her father asks?

a) 6 b) 12 c) 18 d) 24 e) 7

Source: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/enem>

Figure 3.1: Example of ENEM problem classified as in the Combinatorics domain.

Other problems were less easily classified. This was the case, in particular, of some problems in the intersection between Combinatorics and Probability. In the ENEM exam, some

problems involving probability required students to figure out the cardinality of the set of outcomes of a probabilistic event, and the creation of this sample space arguably involves combinatorial ideas. In these cases, I considered that probabilistic concepts played the central role in solving the problem and classified them as non-Combinatorics problems. Figure 3.2 provides an example of a problem involving sample spaces that was classified as a non-Combinatorics problem. This decision was made because the overall structure of the problem is more representative of problems in the domain of probability. The problem explicitly asks “What is the probability” of a given scenario. It thus requires students to have prior knowledge in this domain to be able to solve it.

ENEM 2018 - Question 149, blue version of the exam

A woman has done an ultrasound and discovered that she is pregnant of quadruplets. What is the probability that she will give birth to two boys and two girls?

A) $1/16$ B) $3/16$ C) $1/4$ D) $3/8$ E) $1/2$

Source: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/enem>

Figure 3.2: Example of ENEM problem classified as not in the Combinatorics domain.

As illustrated in the provided examples, the ENEM exam items are closed-ended, with five possible answer choices. On the test date, students randomly receive one of four different exam versions coded by color. The versions vary only in the order in which the exam questions appear within each subject section (e.g., Mathematics, Natural Sciences) - but for any given question, the order of the closed-form possible answers remains the same.

For all quantitative procedures, I focused on the subset of students who completed high school in the same year they took the ENEM exam. This decision was made for two primary reasons. Firstly, this specific subgroup is more closely comparable to the interviewees from the

study's qualitative portion. Secondly, from 2009 to 2015, only final-year high school students responded to the questionnaire item regarding which type of school they attended (public or private), a vital variable for analyzing accessibility in Combinatorics and non-Combinatorics problems. To ensure consistency in data comparison across years, it was thus necessary to examine this particular student subset. My analysis was also limited to students with complete data for all other variables in the study, including race and the series of answer choices given in the Mathematics section of the exam. The resulting subset of students remains substantial, with over 800,000 individuals in each exam year.

Research Question 1

To compare *difficulty* levels between problems in Combinatorics and problems from other domains, I measured difficulty as the rate of incorrect responses for Mathematics problems in the ENEM exam. The first and most straightforward procedure I use for this comparison is a t-test to capture the differences in the rates of incorrect responses between Combinatorics and non-Combinatorics problems (Model 1).

Although this comparison is helpful as an initial descriptive approach, it might be subject to two confounding effects. The first possible confounding variable is the exam cohort. Different years of the exam had different pools of test-takers, who might have varying levels of mathematics preparation. For example, test takers in 2019 might have been better prepared, on average, than their peers who took the exam a decade before, in 2009. If exceptionally or low (or high) achieving cohorts of students coincide with years in which relatively more Combinatorics problems were present in the exam, then my estimates of the difficulty of Combinatorics items might be biased upward (or downward). To control for this potential confounding effect, I include controls for the exam-year cohorts in a multivariable linear regression model (Model 2).

Using the procedure suggested by Hardy (1993), I do this by including year fixed effects (year dummies) in a regression where the dependent variable is the question difficulty level (rate of incorrect response), and the independent variables are (i) a dichotomous variable indicating whether the problem is from Combinatorics or not (my variable of interest), (ii) year fixed effects (control variables). This model aims thus to control for possible differences in achievement attributed to year-specific higher/lower average cohort achievement. I then use t-tests to determine whether the coefficient of interest in the regression - the Combinatorics dummy variable - significantly differs from zero.

A second possible confounding effect is the position of each question within the ENEM exam. Questions at the end of exam booklets have a higher chance of not being answered by students - for example, due to fatigue or time management problems - and have been documented to have higher rates of incorrect responses (Debeer & Janssen, 2013; Kingston & Dorans, 1984). If Combinatorics problems are distributed in a particular way - for example, at the end (or beginning) of the exam booklet, then estimates of difficulty levels may be biased upward (downward) if appropriate covariates are not included. To control for this possible confounding effect, I used a regression model that expands on Model 2 with the inclusion of *item position* in the exam booklet as an additional control variable (Model 3). This variable is intended to capture the relationship between item position in the test and difficulty levels, removing its potential confounding effect on the main variable of interest (the Combinatorics dummy variable).

The three models with increasing levels of controls for possible confounding effects are summarized in Figure 3.3. To make the notation uniform across the models, I also use regression notation for the first model, even though it comprises a simple t-test. This notation is valid

because, as explained by Allison (1999), “When there are only two groups, ANOVA is equivalent to a conventional t-test for a difference in means and, therefore, it's also equivalent to a multiple regression with a single, independent dummy variable” (p. 176). In the notation of Figure 3.3, the coefficient of interest across all three models is thus represented by the same term, β_1 . This coefficient captures the difference between the average difficulty levels of Combinatorics and non-Combinatorics items - first measured without any controls (Model 1), and then with controls for student cohort (Model 2) and also for item position (Model 3).

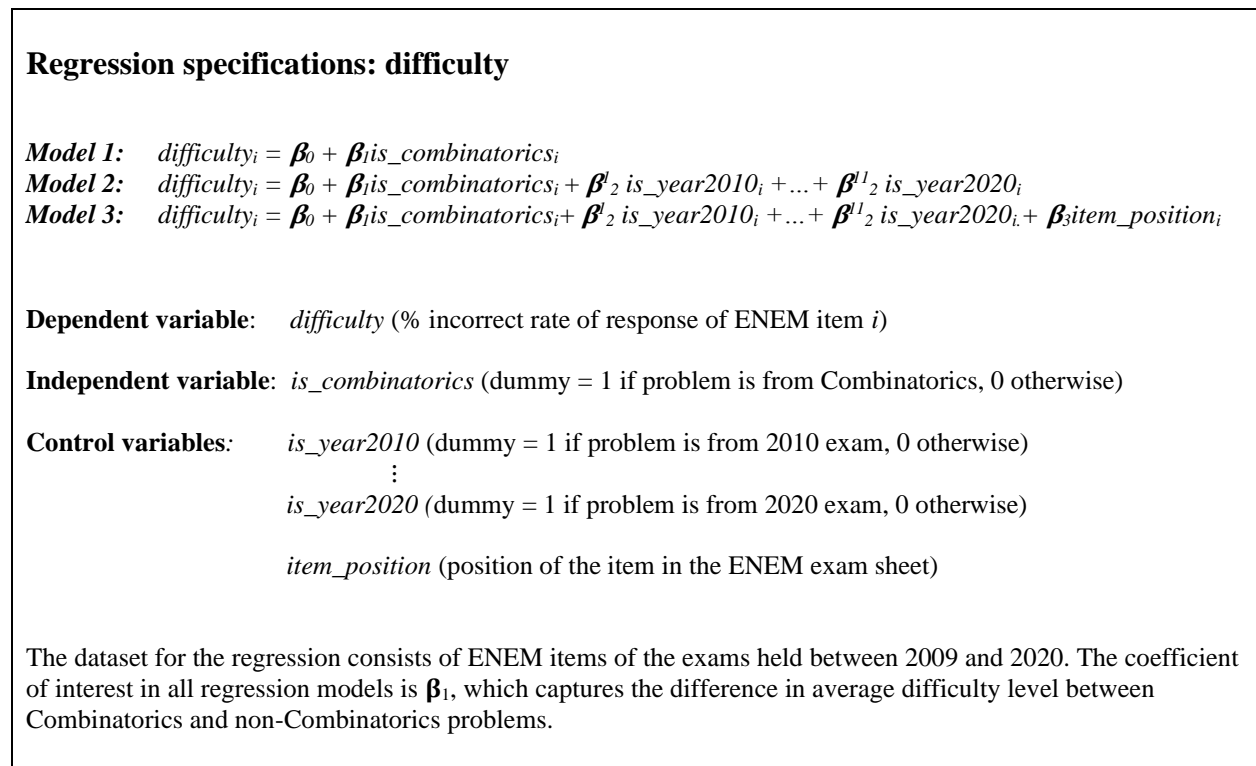


Figure 3.3: Empirical procedure for the first part of Research Question 1 (difficulty).

The second part of Research Question 1 asks about how Combinatorics problems compare to other areas of mathematics (i.e. non-Combinatorics problems) regarding *accessibility*. According to the definition of accessibility proposed in this study, a domain in Mathematics is particularly accessible if, in that domain, students with varied backgrounds can

understand what problems ask and engage in valuable mathematical thinking while trying to solve problems. As English (2005) puts it, accessible domains allow a broad range of students, including those “with minimal content knowledge to work towards a solution” (p. 133). A corollary of this definition is that prior knowledge should have a lower predictive power of student achievement in more accessible domains. Looking at accessibility through the lens of the predictive power of prior knowledge is particularly interesting to this study, as it can lead to testable hypotheses using data from ENEM.

If prior content knowledge is less important in Combinatorics than in other domains, then prior formal mathematics teaching should give students less of an advantage when solving Combinatorics problems than non-Combinatorics problems. One way to test this is to compare well-documented gaps between student populations with different levels of access to quality mathematics instruction in Brazil. One such gap is the private-public school gap. On average, students from private schools in Brazil tend to achieve higher scores in standardized mathematics exams than peers in public schools. The literature documented these gaps at the elementary (Souza et al., 2018), middle (Oliveira et al., 2013), and high school levels (Cavalcanti et al., 2010). Based on this documented gap, the following hypothesis is laid out to test the claim that Combinatorics is a particularly accessible domain in Mathematics:

- *If Combinatorics is more accessible than other domains tested in ENEM, then the private-public test score gaps will be smaller in Combinatorics problems than in non-Combinatorics problems.*

Another documented achievement gap in Brazil is its racial gap. On average, white students tend to study in schools with better resources and score higher on average in standardized tests than black students (Madeira & Rangel, 2014; Becker & Arends-Kuenning,

2020). I thus also investigate the hypothesis that the racial achievement gaps are smaller in Combinatorics than in non-Combinatorics problems:

- *If Combinatorics is more accessible than other domains tested in ENEM, then the white-black test score gaps will be smaller in Combinatorics problems than in non-Combinatorics problems.*

Based on the two hypotheses described above, I used two proxies - the *school gap* and the *racial gap* - to test for differences between the accessibility levels of Combinatorics and non-Combinatorics problems. All confounding variables described in the regression models of item *difficulty* are also potentially relevant in the case of the proxies for accessibility. I thus include the same control variables (*year* and *position* in the test) in the regressions used to investigate *accessibility* (models 1 to 2, figure 3.5). Also, there is now a third confounder: each item's *difficulty level*. This confounding variable is more nuanced and deserves an extended discussion.

An item's difficulty level is a confounding variable to the proxies of accessibility (school gap and racial gap) because an a priori, "mechanical" relationship exists between the two variables. We can explore this relationship by analyzing a few examples. First, suppose that a question in ENEM is so difficult that its difficulty level is 1. For the overall rate of incorrect response to be 1, it must be true that all private school students, as well as all public school students, responded incorrectly (see point *b* in Figure 3.4). An achievement gap of 0 is the only possible compatible scenario - because both groups achieve at the same levels. Another extreme case leads to the same conclusion. If all students answer a problem correctly (i.e. difficulty level is 0), then there must be no difference between the average achievement of each subgroup of students, because all students from *both* groups must have responded to the problem correctly. Thus, in this scenario, the private-public school gap will also be zero (see point *a* in Figure 3.4).

Only for the cases between those two examples (i.e., difficulty levels between 0 and 1) will the proxies of accessibility assume non-zero values. Also, questions with intermediary values of difficulty tend to allow for a larger gap. For example, if the public and private school populations of students are of the same size, then the largest possible gap is associated with a difficulty level of 0.5 and occurs when all private students answer a problem correctly while all public students answer it incorrectly (see point *c* in figure 3.4).

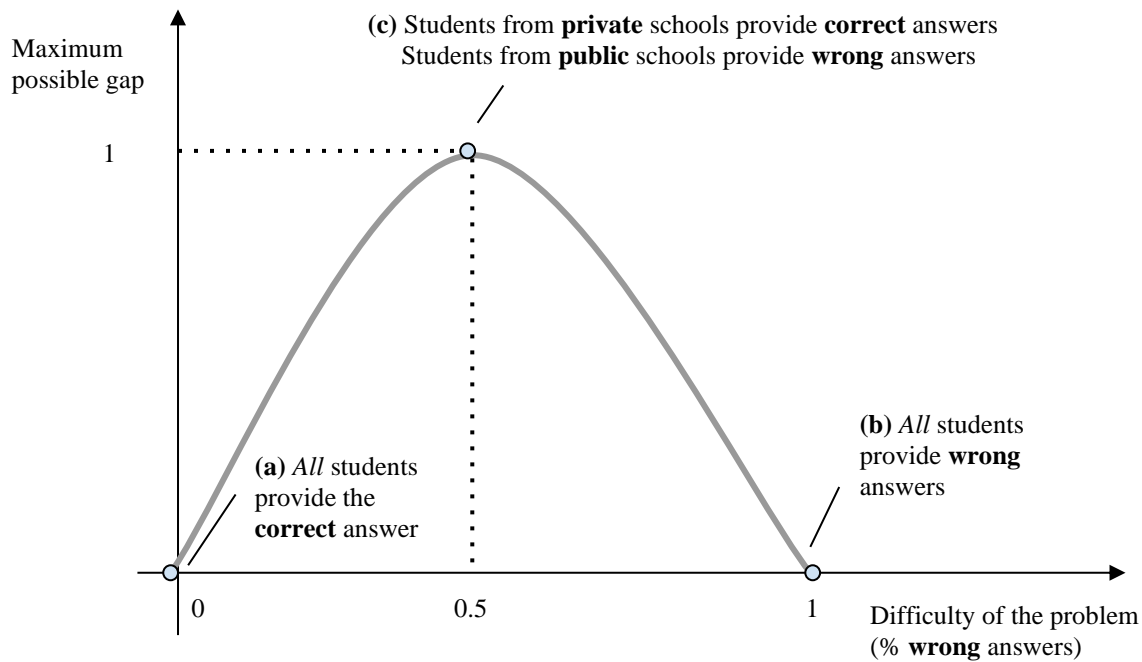


Figure 3.4: Hypothetical maximum school achievement gaps for varied difficulty levels.

To control for this potentially relevant confounding effect of difficulty on the private-public school gap (and racial gap), I resort to two empirical strategies. First, I include difficulty as an additional control variable in the regressions. I use a quadratic model to approximate the relationship between problem difficulty and the private-public school gap discussed above, as shown in models 4a and 4b of Figure 3.5.⁸ Figure 3.5 shows all specifications used to test

⁸ Note that in this model I drop the variable *item_posion*. I made this choice of specification because the rationale for item position to be confounder to the estimates of the school (or racial) gaps is through its influence on *difficulty*,

whether accessibility levels are higher for Combinatorics than non-Combinatorics problems in ENEM.

Regression specifications: accessibility

Model 1a: $school_gap_i = \beta_0 + \beta_1 is_combinatorics_i$

Model 2a: $school_gap_i = \beta_0 + \beta_1 is_combinatorics_i + \beta^l_2 is_year2010_i + \dots + \beta^{l_2}_2 is_year2020_i$

Model 3a: $school_gap_i = \beta_0 + \beta_1 is_combinatorics_i + \beta^l_2 is_year2010_i + \dots + \beta^{l_2}_2 is_year2020_i + \beta_3 item_position_i + \beta_3 (difficulty_i)^2$

Model 4a: $school_gap_i = \beta_0 + \beta_1 is_combinatorics_i + \beta^l_2 is_year2010_i + \dots + \beta^{l_2}_2 is_year2020_i + \beta_3 difficulty_i + \beta_3 (difficulty_i)^2$

Model 1b: $race_gap_i = \beta_0 + \beta_1 is_combinatorics_i$

Model 2b: $race_gap_i = \beta_0 + \beta_1 is_combinatorics_i + \beta^l_2 is_year2010_i + \dots + \beta^{l_2}_2 is_year2020_i$

Model 3a: $race_gap_i = \beta_0 + \beta_1 is_combinatorics_i + \beta^l_2 is_year2010_i + \dots + \beta^{l_2}_2 is_year2020_i + \beta_3 item_position_i + \beta_3 (difficulty_i)^2$

Model 4b: $race_gap_i = \beta_0 + \beta_1 is_combinatorics_i + \beta^l_2 is_year2010_i + \dots + \beta^{l_2}_2 is_year2020_i + \beta_3 difficulty_i + \beta_3 (difficulty_i)^2$

Dependent variable: $school_gap_i$ (difference in the percentage of correct responses for item i between private and public school students)
 $race_gap_i$ (difference in the percentage of correct responses for item i between white and black students)

Independent variable: $is_combinatorics$ (dummy = 1 if problem is from Combinatorics, 0 otherwise)

Control variables: $is_year2010$ (dummy = 1 if problem is from 2010 exam, 0 otherwise)
 \vdots
 $is_year2020$ (dummy = 1 if problem is from 2020 exam, 0 otherwise)

The dataset for the regression consists of ENEM items of the exams held between 2009 and 2020. The coefficient of interest in all regression models is β_1 , which captures the difference in average difficulty level between Combinatorics and non-Combinatorics problems.

Figure 3.5: Empirical procedure for the first part of Research Question 1 (accessibility).

Second, to take extra care about the potential confounding effect of difficulty levels on the estimates of accessibility, in addition to the regressions discussed in the previous paragraph, I estimate accessibility levels using propensity score matching techniques. The motivation to run

which is already included in this model. In any case, I did run the model including both controls (for item position and difficulty) and that did not change the overall results for the coefficients of interest.

these additional estimates is only to provide an alternative - an arguably more conservative - form of controlling the potential confounding effect of difficulty levels on the estimates of accessibility. While regression and matching “are both control strategies” (Angrist and Pischke, 2008, p.69), matching can be interpreted as a more explicit model in how it deals with the control variables. In this type of model, a new sample of Combinatorics and non-Combinatorics items is constructed with the goal of pairing up observations that have similar values in the potentially confounding variables. To do this, the propensity score matching procedure pairs each Combinatorics problem in the database with its non-Combinatorics ‘nearest neighbor’ according to the selected matching variables. I use two matching variables to calculate the minimum-distance neighboring pairs: *item position* in the test and *item difficulty* level. I also impose exact matching at the exam booklet level - in other words, each pair of items (a Combinatorics problem and a non-Combinatorics problem) must belong to the same exam year and version. This imposition guarantees that within the matched pairs of problems, the difficulty levels are calculated based on the same population of test takers and the same test version, thus rendering them more directly comparable. All three regression specifications depicted in Figure 3.5 are then rerun with the new, matched sample of ENEM problems.⁹

⁹ In principle, since the matching procedure creates comparable samples of Combinatorics and non-Combinatorics problems, we could just do a direct t-test to compare the accessibility levels of the two types of problems, dispensing the covariates of the multiple variable regression specifications. Adding the matching variables as controls in a multivariable regression is a more conservative practice, though, as it allows for the additional control on the remaining imbalances in covariate levels between the two groups (Combinatorics and non-Combinatorics problems). If the matching procedure is sound (i.e. the two samples are sufficiently similar in regard to the confounding variable), we should expect the multivariable regression estimates to be fairly stable across the four specification models. The matching of the year-fixed effects controls, on the other hand, is exact, in the sense that each Combinatorics problem is paired with a non-Combinatorics problem in the exact same year. The addition of the year fixed-effect covariates (Model 1 vs Model 2) should thus have no effect at all on the estimates of the coefficient of interest. So in this case, maintaining the Model 2 is just convenient to make the final tables of the matching procedure more easily comparable the previous specification.

The hypotheses tested to address Research Question 1 are represented in Figure 3.6. The empirical procedures employed test only whether the levels of difficulty and accessibility are statistically different among Combinatorics problems in comparison to non-Combinatorics problems. No investigation has been proposed so far in regard to explanatory factors for these eventual differences (shaded area).

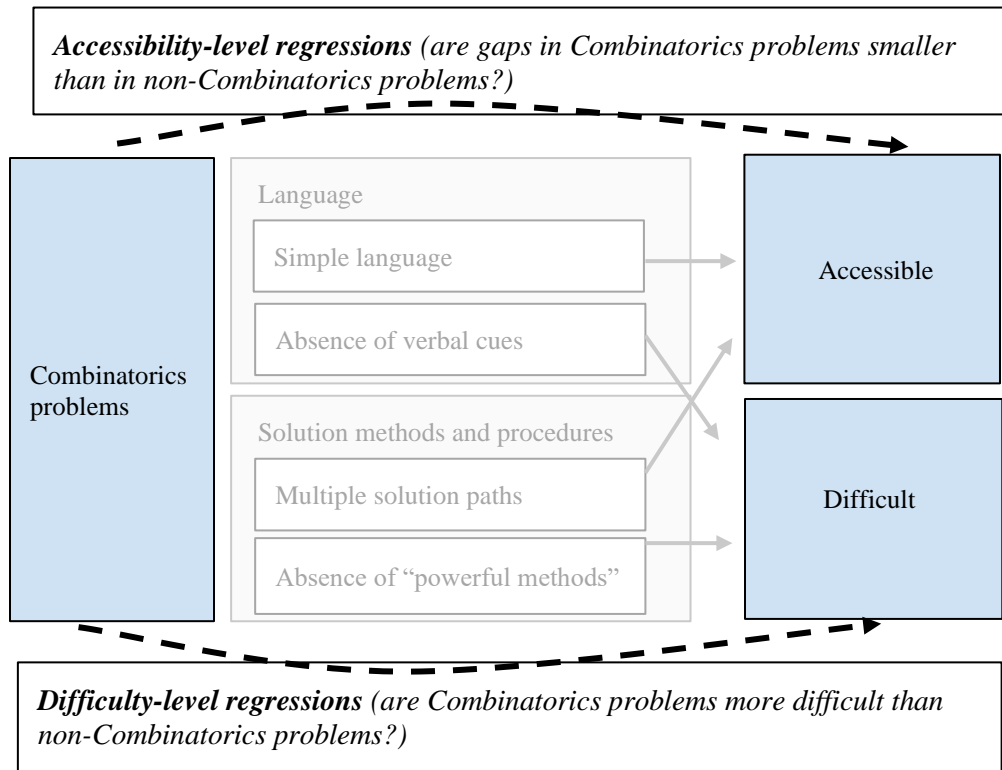


Figure 3.6: Hypotheses tested to address Research Question 1.

Research Question 2

The second proposed research question focuses on predictors of difficulty and accessibility levels across ENEM problems. This research question is also approached from a quantitative perspective. First, I quantified two characteristics of the ENEM Combinatorics problems that the literature suggests being possible determinants of accessibility and difficulty. I

then perform statistical tests to check whether any of these characteristics are predictors of item difficulty and accessibility levels.

The first quantifiable characteristic I analyze is *item length*. As pointed out in Daroczy et al.'s (2015) literature review, prior research has indicated that mathematics problems with longer statements tend to present higher difficulty levels. Daroczy et al. (2015) also highlight that many other indicators of linguistic complexity in mathematics problems are associated with difficulty levels, such as grammatical complexity and average length of sentences. Among these characteristics, though, I focus only on item length due to convenience. That is, length was the only variable that could be straightforwardly quantified within the time and resources available for this study. The quantification of item length was indeed straightforward. The text of each problem in ENEM was copied to a spreadsheet using *Google Sheets*, and an automated function was used to compute their word counts. I included only the words in the questions' written statements - all words that were part of tables, answer choices, drawings, or diagrams were thus excluded from the word count.

As discussed in the previous chapter, the mathematics education literature indicates that many Combinatorics problems can be written without mathematics notation and symbols, and that this may be one of the factors associated with accessibility in this domain (Lockwood et al. 2020, Sandefur et al., 2022). To investigate this claim, I quantified the presence of mathematical notation in all ENEM problems. This procedure is intrinsically subjective, involving the researcher's judgment of the degree of complexity of mathematical symbols. To increase the reliability of the classification used, I conferred with two experienced high school teachers to develop a list of symbols that would be classified as complex. Our consensus led to the following list of complex symbols:

- Algebraic symbols used to represent variables in a mathematical expression (e.g., letters representing unknown quantities)
- Greek letters used in any context (including pi)
- Logs
- Symbols that represent combinatorial operations (binomial coefficients, arrangements, factorials)
- Radians
- Non-integer exponents
- Square roots
- Matrices and determinants

To support the classification of dubious terms, the two high school teachers and I also created a list of symbols that would not be considered complex in the coding process:

- Natural exponents (including units of volume measurement)
- Fractions
- Percentages
- Measurement units, in general,
- Currency symbols

As an attempt to reduce the level of subjectivity in the classification, we decided to use only symbols (not the mathematical meaning of words) in our list. In favor of objectivity, I coded the *presence of mathematical notation and symbols* as a binary variable. That is, each problem was coded as 1 if it included any symbol in the list and as 0 otherwise. Problems with one symbol and problems with many symbols are thus indistinguishable in the coded data. While information is lost in this process (as compared, for example, to coding how many times a symbol in the list is present or how many symbols are present), it made the coding process less time-consuming and more feasible in the available time for this research to be carried out.

To answer Research Question 2, I add these two quantified characteristics (item length and item notation) as covariates in the regression equations used in Research Question 1. The results of the new specification models are analyzed in two ways. First, I check whether the inclusion of the item-specific variables “explains away” eventual differences in difficulty and accessibility between Combinatorics and non-Combinatorics problems (i.e. once the additional

variables are taken into account, do difficulty and accessibility levels become statistically similar across Combinatorics and non-Combinatorics problems?). Second, I analyze the sign and significance of each variable's coefficients in the regressions (i.e., is item length a predictor of difficulty and accessibility levels in ENEM problems? Is the presence of notation a predictor?). The empirical models used to address Research Question 2 are described in Figure 3.7. All models are run as regressions, in which the coefficients of interest are β_I (combinatorics dummy) β_L (item length) and β_N (notation).

Additionally, I test whether the two linguistic factors studied (item length and item notation) present differential patterns among Combinatorics and non-Combinatorics problems. To do that, I separately reproduce the empirical procedure summarized in Figure 3.7 across the subsamples of Combinatorics and non-Combinatorics problems.¹⁰ If the predictions that Combinatorics accessibility levels result (at least in part) from the simplicity in which Combinatorics problems are stated, we should find negative correlations between *accessibility* levels and each of the two variables (length and notation). The literature also suggests that statement simplicity does not necessarily make Combinatorics problems less difficult. Consequently, I predict that this characteristic should not be a strong explanatory factor for difficulty levels within the subsample of Combinatorics problems.

¹⁰ I drop the *is_combinatorics* dummy variable in these regressions since they are run on subsamples comprised only of Combinatorics problems (or only of non-Combinatorics problems).

Regression specifications: Research Question 2

$$\text{Model 1: } \text{difficulty}_i = \beta_0 + \beta_I \text{is_combinatorics}_i + \beta'_2 \text{is_year2010}_i + \dots + \beta^{I,2} \text{is_year2020}_i + \beta_3 \text{item_position}_i \\ + \beta_L \text{length} + \beta_N \text{notation}$$

$$\text{Model 2: } \text{school_gap}_i = \beta_0 + \beta_I \text{is_combinatorics}_i + \beta'_2 \text{is_year2010}_i + \dots + \beta^{I,2} \text{is_year2020}_i \\ + \beta_3 \text{difficulty}_i + \beta_3 (\text{difficulty}_i)^2 \\ + \beta_L \text{length} + \beta_N \text{notation}$$

$$\text{Model 3: } \text{race_gap}_i = \beta_0 + \beta_I \text{is_combinatorics}_i + \beta'_2 \text{is_year2010}_i + \dots + \beta^{I,2} \text{is_year2020}_i \\ + \beta_3 \text{difficulty}_i + \beta_3 (\text{difficulty}_i)^2 \\ + \beta_L \text{length} + \beta_N \text{notation}$$

Dependent variables: *difficulty* (rate of incorrect responses for item *i*)
school_gap_i (difference in the percentage of correct responses for item *i* between private and public school students)
race_gap_i (difference in the percentage of correct responses for item *i* between white and black students)

Independent variables: *length* (item word count)
notation (dummy = 1 if the problem contains mathematical notation, 0 otherwise)

Control variables: *item_position* (position of the item in the ENEM exam sheet)
item_difficulty (overall percentage of correct responses for item *i*)
item_difficulty² (overall percentage of correct responses for item *i*, squared)

The dataset for the regression consists of ENEM Combinatorics items of the exams held between 2009 and 2020. The coefficient of interest in regression models is β_L and β_N , which capture the linear associations between item's characteristics (*length* and *notation*, respectively) and their levels of difficulty and accessibility, and β_I , which captures the difference in average levels of difficulty (or school or racial gaps) between Combinatorics and non-Combinatorics problems.

Figure 3.7: Empirical procedure for the first part of Research Question 2

The analysis proposed above to answer Research Question 2 addresses only a modest part of the hypothesis raised in the previous chapter, as summarized in Figure 3.8. An analysis of the remaining hypothesis is conducted in the qualitative analysis employed to address Research Question 3.

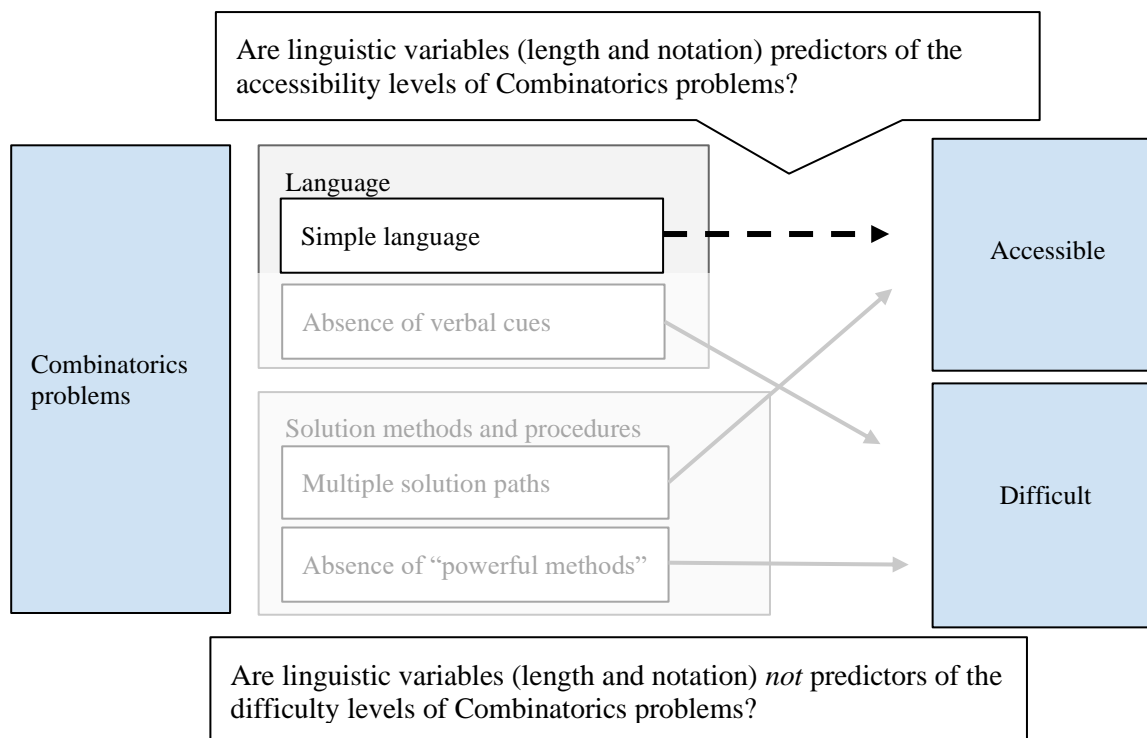


Figure 3.8: Hypotheses tested to address Research Question 2.

Qualitative Data and Research Question 3

The third research question asks *how* students understand and engage with Combinatorics problems, as well as what prevents them from arriving at correct responses. To answer this question, I interviewed Brazilian students using a task-based interview design (Maher & Sigley, 2020). I purposefully selected eight Combinatorics and four non-Combinatorics problems from the ENEM exam according to criteria that will be discussed further in this chapter. I interviewed each student individually to gather information on how they interact with this subsample of problems. Data were collected through a demographic questionnaire, video recordings and transcripts of the interviews, field notes, and pictures of students' written work.

Research subjects

Because this study is mainly interested in understanding how Combinatorics relates to issues of equity in access to mathematics, I argue that the most appropriate participant profile for my task-based interviews would be that of students with disadvantaged school-mathematics backgrounds. Analyzing how students with disadvantaged backgrounds interact with Combinatorics problems can show how prior content knowledge (or the lack thereof) interferes with how students understand and attempt to solve problems in this domain.

With this participant profile in mind, I planned to interview 15 students from a philanthropic preparatory course for the ENEM exam, located in a large metropolitan area in Brazil. As a philanthropic entity, the preparatory course accepts only students from low-income families who have studied in public high schools. I also restricted the pool of participants to students who had taken their last year of high school in 2021, the year before the interviews were held (2022). This restriction was imposed to increase the compatibility between the ENEM and the interview samples. As compensation for their study time lost during the interviews, I offered students a bookstore voucher of R\$ 150 (roughly \$30) and advised them to buy books that could help them study for the ENEM exam.

The management of the philanthropic preparatory course was supportive of student participation in the research. After signing the IRB form, a coordinator in the philanthropic course sent out research flyers to their student lists using a messaging phone app. While 17 students responded to the flyer manifesting their intentions to participate in the research, only 9 responded to my emails and showed up for the interviews on their scheduled dates. Because this was still an insufficient number of participants (my initial goal was 15), I contacted the 9 interviewed students and asked them to send the research flyer to former high school colleagues.

This “snowball” strategy led to 9 additional students responding to the flyer, but only 4 ended up showing up on their scheduled interview dates. In total, I was thus able to interview 13 students - 9 from the philanthropic preparatory course and 4 students who were their friends in high school.

Interview setup

The interviews were held individually via Zoom video call in the month of September 2022. Holding the interviews remotely had three benefits. First, it prevented the interviewer and interviewee from being exposed to Covid at a time in which the pandemic still held high numbers in Brazil. Second, it simplified scheduling by overcoming the need to find a physical place to meet. Third, recording a Zoom meeting is more straightforward and reliable than setting up a camera and microphones in an interview room, thus making the documentation of the interview simpler and less prone to failure.

Although the video call arrangement simplified some aspects of the data collection, there were also challenges associated with using this form of communication. First, students without access to broadband from their houses were, in principle, precluded from participating. To mitigate this problem, the participant recruitment flyer informed students that they could do the interview from any quiet location with broadband - not necessarily their homes. In fact, all except one interviewee connected from their households. A second problem was that in some interviews, students were interrupted briefly by their family members or by some household event (e.g., someone knocking on the door). These interruptions were short and did not bring significant problems to the interviews in all except one case. The problematic case involved one female student being asked by her family to help take care of a younger sibling. She asked to interrupt her interview after responding to half of the questions. I decided to exclude this interview from my analysis because this was the only partially complete interview.

Interview protocol

The interviews were intended to uncover aspects of students' interactions with ENEM problems that could shed light on both access and difficulty of Combinatorics and non-Combinatorics problems. With that goal in mind, I designed an interview protocol that prompted students to explain their understanding of the problems, as well as the mathematical thinking they used in attempting to solve them. To develop the protocol details, I did two rounds of pilot interviews (November 2021 and May 2022), each of them with two subjects. I describe an overview of the interview process in the next paragraphs and provide the detailed protocol in Appendix A.

The interviews were set to last 90 minutes and began with an opening statement in which I briefly explained the purpose of the interview and the types of questions I would ask. This was also a time for students to ask any questions they had about their participation. I learned from the pilot interviews that my initial communication should inform students about what they should expect my behavior to be during the interview. In particular, I was advised to highlight to the interviewees that I would always ask for explanations about their thinking, regardless of my judgment of whether their answers were correct or incorrect. To help students feel comfortable sharing their mathematical ideas, I explained that my research interest was in how people interact with different problems in mathematics, as opposed to just checking whether they were getting answers correct. I told them that by sharing their thinking, they were contributing to the research even if they answered the problems wrongly. After this introduction, I shared my screen and went by each of the 12 problems on the list.

For each problem, I first asked students to read the question on the shared screen (either aloud or silently). With this first reading, I asked them to try to understand what the problem was

asking but not yet to attempt to solve it. When they were finished reading the problem, I asked them whether they felt that they understood what the problem was asking, trying to emphasize that one can understand a problem event without knowing exactly how to find the answer. This question was intended to provide information on students' own perceptions about whether they understood what each problem asked.

In case they answered positively (i.e. they felt like they understood what the problem was asking), I told them to explain what the problem was asking *in their own words*. With this question, I intended to gather information about the mathematical soundness of their understanding of the problem statement. After they rephrased the problem, I also asked them to explain the meaning of specific mathematical vocabulary that I pre-identified as possibly challenging technical words or symbols. This list of technical terms was created in consultation with two experts - Brazilian high school mathematics teachers - to whom I showed all 12 questions and asked them to point out words or symbols that students may find difficult to interpret. By asking students to explain those terms, I intended to find out whether they were prerequisite knowledge barriers preventing students from understanding the problems.

Finally, I told students to solve the problem and allowed them silent time to work on it. After the student arrived at an answer or gave up trying, I asked them to explain their thinking. For students who gave up or arrived at an answer different from the available choices, I insisted that I would like to understand how they tried to solve the problem. That was, in general, the most interactive exchange part of the interviews. I would sometimes interrupt students to ask for clarification or probe them to tell me if they also tried to solve the problem in different ways, even if unsuccessfully. If students took a too long time on a specific problem, I would ask them to share with me their partial work and to pick one of the five possible answers as their "best

current guess”. I set the reference time for how long they could spend on a problem as 7 minutes and shared this information with them at the beginning of the interview so they would not be surprised by my suggestions to move to the next problem. I did not follow this reference tightly, though, because in some cases students went quickly through some of the questions, allowing for more time to be spent on other problems.

Data sources

The transcripts of the recordings of the 13 interviews constitute the major data source of the qualitative portion of this study. They are complemented by three additional data sources: questionnaires, field notes, and pictures of student work. This varied documentation of students’ interactions with the interview problems allowed me to triangulate information (Creswell & Poth, 2017; Mathison, 1988), looking at the additional sources when a need for additional details emerged in the analysis of the interview transcripts.

The approach to creating transcripts was twofold. First, I used the automated transcription service *Transkriptor*. I then listened to the audio recording of each interview and edited the initial transcript version provided by the software to correct for mistakes. After all individual interviews were transcribed, I gathered all of them in a single text file, which was the main document used in my coding analysis.

Students also responded to a short questionnaire before the date of their interviews. The goal of the questionnaire was to gather demographic information about the interviewees, including the school and year in which they graduated, age, gender, and ethnicity. I also included two open-ended questions about their experience with mathematics in high school. The first question asked students to describe a positive experience they had with mathematics in high school, and the second asked them to describe a negative experience with the subject.

During each interview, I took field notes in a digital file. My main goal with notetaking was to document my thoughts and interpretations about students' interactions with each problem. I also took notes about any interference to the interviews, such as the planned breaks and unplanned interruptions. As a complementary source of information on student thinking, I asked them to send me pictures of all the annotations they made in trying to solve the problems during the interview. These images were sent to me by 12 of the 13 students.

Selection of problems

To create the list of questions from ENEM that would be used in the interviews, I resorted to a purposeful selection of problems of varied levels of difficulty and accessibility. The rationale for the selection was to include the whole spectrum of combinations of high/low levels of difficulty and accessibility. At this point in the research process, I already had estimated numerical measures for the difficulty as well as the *proxies* for accessibility of each Combinatorics problem, which were used as the first step in the problem selection. After selecting a shortlist of questions with the highest and lowest levels of difficulty and the highest and lowest levels of accessibility (proxy), I consulted with two mathematics high school teachers and showed them the problems. From this shortlist, we selected problems that were varied in terms of their combinatorial content (e.g. arrangements, permutations), their form (e.g. presence or absence of mathematical notation), and expected levels of accessibility, according to the experts' perspectives. After the Combinatorics problems were selected, I selected four "non-combinatorics peer questions" that were the closest in difficulty level (as measured based on the ENEM data) to four of the selected Combinatorics questions. Based on this logic, 12 problems were selected for the questionnaire - 8 of them from Combinatorics and 4 from other domains. The full list of interview problems is reproduced in Appendix B.

Overview of the coding process

To analyze the transcribed data, I first created a single text file organized as a large table. Transcripts of different students were assigned to each row in the table, and problems were distributed over columns. This organization allowed the data to be read either student-by-student or problem-by-problem, according to need. The coding process used in this study's qualitative analysis comprised two phases. The first round of coding was intended to provide a summary of the overall levels of accessibility and difficulty of the different problems in the problem set. In this first phase, I coded students' levels of understanding of what problems asked, their engagement with mathematics processes, and the correctness of their responses. The second phase in the coding process was a thematic analysis aimed at uncovering phenomena that could shed light on the third research question. While I started this analysis with an initial set of pre-defined themes, I remained open to identifying new, unexpected phenomena emergent in the data. The first and second coding phases are described in more detail in the next paragraphs.

First phase of coding - accessibility and difficulty across problems

The first phase in the qualitative analysis process was the coding and tabulation of data for indicators of accessibility and difficulty. This analysis was done for one problem at a time, to allow for a consistent assessment across the responses to any given problem (i.e. all transcripts of Problem 1 were firstly coded, then all transcripts for Problem 2, and so forth). The criteria used in this initial coding is reproduced in Table 3.1.

Table 3.1: Criteria for the first round of coding.

Theme	Criteria
Accessibility - Student self-perception about whether they understood what the problem asked.	<p>Yes. The student stated that they understood what the problem asked.</p> <p>Partial. The student stated that they partially understood what the problem asked.</p> <p>No. The student stated that they did not understand what the problem asked.</p>
Accessibility - Researcher rating of whether the student understood what the problem asked.	<p>Yes. The student rephrased the problem in a way that would allow a listener with knowledge about the mathematics content used in the problem to understand it. Minor numerical information may have been omitted.</p> <p>Partial. The student rephrased the problem in a way that the listener can partially understand. No conceptual misguidance is given in the statement, but significant information is missing.</p> <p>No. The student rephrased the problem in a way that would mislead a listener about what is asked in the problem, or provides insufficient information for the listener to understand what is being asked.</p>
Accessibility - Mathematical sensemaking in solution attempt	<p>Conceptual engagement. The student engaged in conceptual mathematical sensemaking while attempting to solve the problem (even if arriving at an incorrect response).</p> <p>Procedural engagement. The student engaged in an exclusively procedural attempt to solve the problem. No evidence is present about conceptual sensemaking.</p> <p>No engagement. The student does not engage with mathematics while trying to solve the problem. The student provides a guess without mathematical justification, or refuses to work on the problem.</p>
Difficulty - Correctness of student response choice.	<p>Correct. The student chooses the right answer.</p> <p>Incorrect. The student chooses a wrong answer or refuses to answer the problem.</p>
Accessibility/Difficulty - Correctness of student’s explanation of their response choice.	<p>Correct. The student explains a solution method that is mathematically coherent and is linked to a correct answer; or the student explains in detail a solution method that can adequately be used to solve the problem even if incurring in a numerical error (but not in a conceptual error).</p> <p>Incorrect. The student provides an explanation that is inconsistent with the correct response to the problem due to a conceptual flaw; or the student explicitly guesses the answer without a mathematical justification for their choice; or the student does not provide an explanation for their answer choice.</p>

The coding of students’ levels of understanding of what each problem asked was directly related to questions I posed to them, as documented in the Interview Protocol. For example, I asked students “Do you think you understood what the problem is asking?” and “Please tell me in your own words: what is this problem asking?”. The data used in this coding process was thus

mostly comprised of interview transcripts, with eventual consultations to interview notes when complementary evidence was needed.

The coding of student engagement with mathematics processes was based both on the interview transcripts and on the pictures of student's written works. I initiated the coding process by looking for evidence of student procedural and conceptual engagement with mathematics in the interview transcripts. I then looked for additional evidence in students' written works.

The coding of the correctness of students' answer choices was the most straightforward among the themes in the initial coding round. Responses were coded based on students' verbal statements about their final answer choices.

Coding the appropriateness of students' explanations of their answer choices was a more nuanced, high-inference process. Reading the transcripts, one problem at a time, I first identified whether each student verbally explained their solution attempts to solve the problems and whether these solutions included evidence of mathematical reasoning that was coherent with solving the problem correctly. I then re-read these instantiations with the support of each student's written work to decide on the final coding.

To support the subsequent, more complex thematic analysis of student's interactions with the interview problems, the final codes of the first round of analysis were documented in the text file containing the transcripts.

Second phase of coding - thematic analysis

The second round of coding of the interview transcripts was initiated using broad themes based on the hypotheses derived from the literature review in the previous chapter. While the quantitative approach employed in addressing Research Questions 1 and 2 addressed a limited scope within these hypotheses, the qualitative analysis employed in Research Question 3 permits,

in principle, the exploration of all connections thereby highlighted. To do that, I searched for possible evidence in students' interview transcripts that could help illuminate the third research question by exploring the four connections discussed in this study's literature review:

- Simple language as a factor related to accessibility
- Multiple methods as a factor related to accessibility
- Absence of verbal cues as a factor related to difficulty
- Absence of powerful methods as a factor associated with difficulty

Following the general guidelines of grounded theory, I also allowed for new coding themes to be developed along the data analysis (Charmaz, 2014; Creswell & Poth, 2019). This strategy is well aligned with the open-ended nature of the research questions addressed by this investigation (*What are some of the ways in which students access problems in Combinatorics? What are some of the ways in which students fail to provide correct answers to problems in Combinatorics? What is the relationship between difficulty and accessibility in students' interactions with Combinatorics problems?*). Finally, in general, and as much as possible, I aimed to take advantage of the pairs of Combinatorics/non-Combinatorics problems to compare observed phenomena across mathematical domains. Whenever needed, I consulted complementary data sources (questionnaires, field notes, student work) to triangulate information.

My coding process for this phase comprised the following steps. First, I read the text files repeatedly, looking for possible phenomena that could shed light on each of those initial themes and adding comments to the data to document these phenomena. I then documented brief descriptions of these phenomena and decided on which of them would be further investigated. My selection prioritized the inclusion of at least one theme related to each of the four

connections listed above. After selecting these phenomena, which are described in detail in the next chapter, I re-read the data to identify additional student interactions in which they were empirically supported.

Chapter IV: Results

This chapter reports on the results of this study. As in the previous chapter, I start with an overview of the data and then follow the order of my proposed research questions to present the results.

Data overview

I identified 28 Combinatorics items across the 12 years (14 exams) of ENEM and 600 non-Combinatorics mathematics items. Table 4.1 shows the number of Combinatorics and non-Combinatorics problems by year of ENEM. In each year, there was at least one Combinatorics problem among the 45 items that comprise the Mathematics section of the test.

Table 4.1. Combinatorics and non-Combinatorics items used in the study.

year	Combinatorics problems	Non-Combinatorics problems	Total
2009	1	44	45
2010	1	44	45
2011	2	43	45
2012	3	42	45
2013	2	43	45
2014	1	44	45
2015	2	43	45
2016	2	43	45
2016 (2 nd offering)	1	44	45
2017	4	41	45
2018	1	43	44
2019	2	43	45
2020	3	41	44
2020 (2 nd offering)	3	42	45
Total	28	600	628

The subpopulation of students whose data are used to calculate the difficulty and accessibility levels is shown in Table 4.2. As explained in the previous chapter, I use the

subpopulation of students who were graduating from high school in the year in which they took the exam. The number of students is over 800,000 for any year in the database.

Table 4.2. Descriptive statistics of subpopulation ENEM test-takers of the study.

	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Race												
<i>Undeclared</i>	2.7%	3.5%	2.6%	1.7%	1.5%	1.4%	1.6%	1.7%	2.0%	2.0%	2.1%	2.1%
<i>White</i>	44.4%	48.3%	46.5%	46.7%	44.6%	43.1%	41.3%	40.2%	39.3%	38.9%	39.7%	42.0%
<i>Black</i>	9.6%	8.8%	9.8%	9.6%	10.2%	10.5%	11.0%	11.5%	11.2%	10.9%	11.2%	10.4%
<i>Parada</i>	39.3%	36.7%	38.3%	39.2%	40.9%	42.4%	43.4%	43.9%	44.6%	45.2%	44.1%	42.7%
<i>Yellow</i>	3.2%	2.0%	2.2%	2.2%	2.2%	2.1%	2.2%	2.2%	2.3%	2.3%	2.3%	2.2%
<i>Indigenous</i>	0.8%	0.7%	0.6%	0.6%	0.6%	0.6%	0.5%	0.6%	0.6%	0.6%	0.6%	0.5%
Graduation school												
<i>Public</i>	78.0%	78.3%	79.1%	77.6%	78.3%	79.1%	79.5%	80.3%	80.1%	78.2%	83.0%	80.0%
<i>Private</i>	22.0%	21.7%	20.9%	22.4%	21.7%	20.9%	20.5%	19.7%	19.9%	21.8%	17.0%	20.0%
Math Score												
<i>Mean</i>	499	505	517	509	510	473	468	490	519	535	523	521
<i>(SD)</i>	(99.2)	(113)	(117)	(121)	(104)	(105)	(107)	(103)	(105)	(103)	(109)	(117)
<i>n</i>	830,175	114,039	1,243,383	1,254,405	1,357,993	1,452,364	1,425,320	1,523,113	1,384,158	1,090,539	1,218,984	815,663

Research Question 1

Research Question 1 asks “How do Combinatorics problems compare to problems in other domains in terms of difficulty and accessibility?” I show histograms in Figure 4.1 to provide an initial visual representation of how difficulty levels vary across Combinatorics and non-Combinatorics problems. The histogram conveys the overall pattern that will remain consistent throughout the regression models - on average, Combinatorics problems are more difficult (i.e. present higher incorrect response rates) than problems from other domains in the ENEM examination. The distribution of difficulty levels across Combinatorics problems is more concentrated at higher levels and presents a higher mean than non-Combinatorics problems.

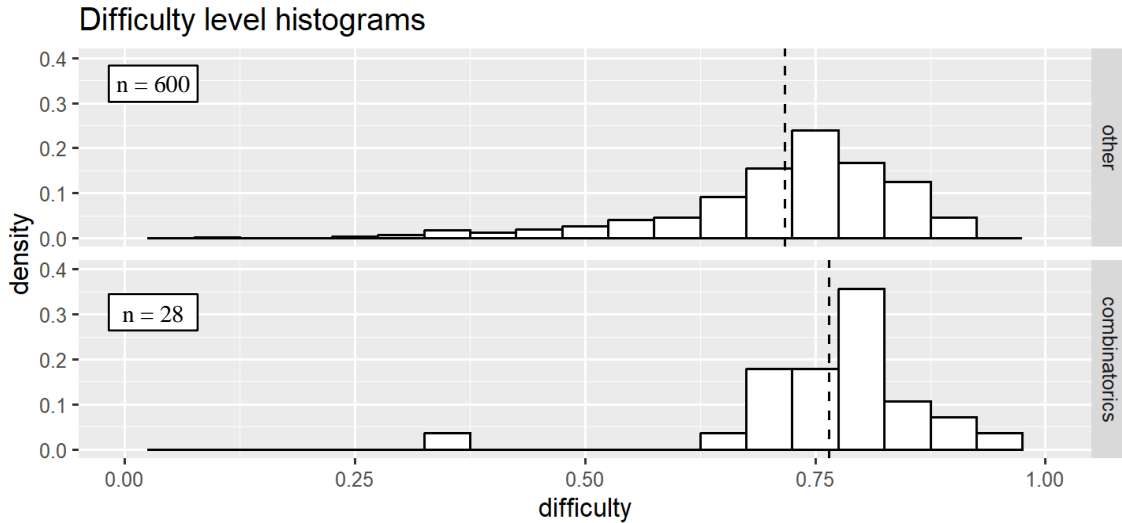


Figure 4.1. Difficulty level distributions of Combinatorics and non-Combinatorics problems in ENEM (2009-2020, grey test version).

Table 4.3 shows the results of the four regression models with an increasing number of controls for potentially confounding variables. All regressions use only the *grey* test version assigned to students randomly on the test date (recall that the questions are the same across the versions, but the order in which they appear on the exam sheet varies). The results of using other test versions are all consistent and lead to the same conclusions and are thus not reported here. Although it would be possible to use the four versions of the test together in the regression models, that would be an arguably less conservative estimation procedure because this procedure would overestimate the item data sample size. Using all test versions, the number of observations in the regressions would be four-fold the original number of different items (i.e. each item would appear four times in the database - once for each test version). Even though there are statistical methods designed to address the issue of repeated observations (e.g. cluster-robust standard errors, Angrist & Pischke, 2008), using a simple test version has the benefit of rendering the results more easily interpretable in regard to sample sizes. By using only the *grey* test version,

sample sizes in regression tables present in this chapter shown in thus adequately represent the number of existing different ENEM problems.

Table 4.3. Regressions of difficulty levels (dependent variable) on a Combinatorics dummy and covariates (ENEM 2009-2020, grey test version).

	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>
<i>(Intercept)</i>	0.717*** (0.005)	0.749*** (0.018)	0.443*** (0.061)
<i>combinatorics</i>	0.048* (0.024)	0.052* (0.024)	0.051* (0.023)
<i>item position</i>			0.002*** (0.000)
<i>year fixed effects</i>	<i>no</i>	<i>yes</i>	<i>yes</i>
<i>R-squared</i>	0.006	0.064	0.105
<i>N (items)</i>	628	628	628

Significance: *** = $p < 0.001$; ** = $p < 0.01$; * = $p < 0.05$. Standard deviations are shown in parenthesis, below each coefficient.

In all three regression models, the coefficient of the dummy variable that captures the difference in difficulty between Combinatorics and non-Combinatorics problems is positive, indicating that Combinatorics problems are more difficult, on average, than non-Combinatorics problems. The coefficients are significant at the 5% level in all models, and the point estimates are fairly constant across the three specifications, indicating that the potential confounders (student cohort and position of the items within the exam) do not play a large role in explaining why Combinatorics problems are more difficult than non-Combinatorics problems. In other words, the control variables can be interpreted as orthogonal to our variable of interest. Nevertheless, the coefficient of the covariable “position” is significant at the 0.1% level, along with some of the coefficients of the year-specific dummies (years 2010, 2011, and 2012). This indicates that such covariates do play a role in predicting questions’ difficulty levels and should

arguably be taken into account for more conservative estimates of our coefficient of interest (the “combinatorics” dummy variable). Considering that, the final and most conservative estimate of the difference in difficulty levels (Model 3) indicates that Combinatorics problems have, on average, a 5.1% higher incorrect response rate than non-Combinatorics problems.

The second part of research question 1 compares the *accessibility* levels of Combinatorics and non-Combinatorics problems. To introduce the results to this question, I also start with a visual representation provided in histograms (Figure 4.2). The first proxy variable I use as an indicator of accessibility is the private-public school gap. It is important to notice that the proxy has an inverted direction regarding our phenomenon of interest - a small gap means a more accessible problem, in which private school students hold less advantage over public school students. As shown in Figure 4.2, the distribution of public-private achievement gaps is concentrated at lower levels for Combinatorics than for non-Combinatorics problems. The same pattern appears in the histograms of the white-black achievement gap, the second proxy I use for accessibility. This gap is also smaller, on average, across Combinatorics problems when compared to non-Combinatorics problems, indicating that Combinatorics problems are more accessible in general.

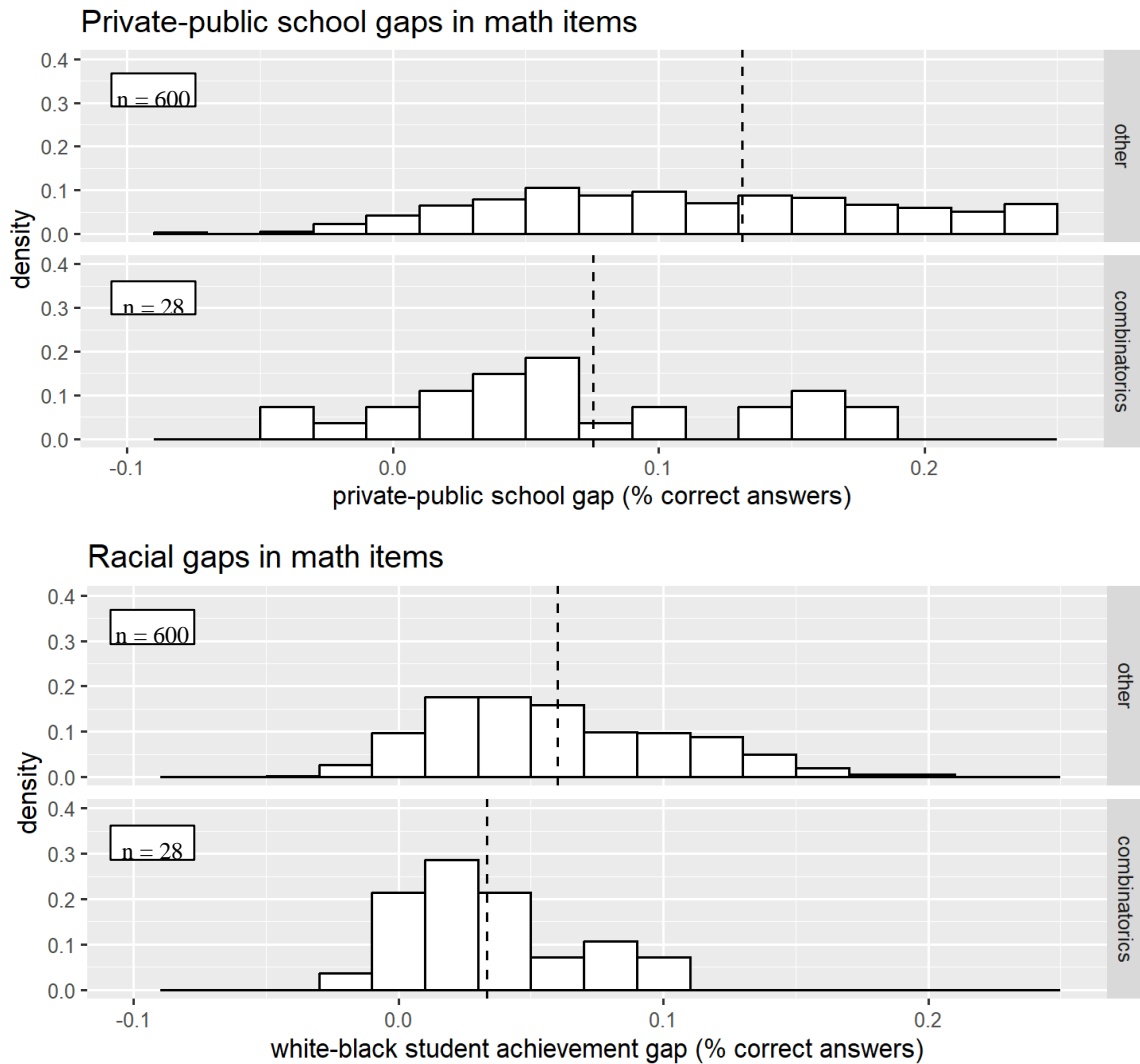


Figure 4.2. School and racial gap distributions of Combinatorics and non-Combinatorics problems in ENEM (2009-2020, grey test version).

As was the case in the analysis of difficulty levels, the regression models of the accessibility proxies of the private-public gap (Table 4.4) and the white-black achievement gap (Table 4.5) confirm the initial evidence presented in the histograms. All models are significant at the 5% level and show effects in the direction predicted by the literature - accessibility is higher in Combinatorics problems, as captured by negative coefficients for the Combinatorics dummies

(smaller school and racial gaps). But differently from the case of the regressions of difficulty level, one control variable seems particularly relevant as a confounder of the association between the pertaining of a problem to the domain of Combinatorics and its accessibility level. As we include controls for item *difficulty*, the point estimates of the gap differences reduce by a sizable magnitude. The *combinatorics* coefficient for the school gap changes from -5.9% in Model 3 to -3.7% in Model 4 and for the racial gap change from -2.9% in Model 3 to -1.5% in Model 4. These results are in line with the discussion made in the previous chapter, which highlighted an intrinsic relationship between difficulty and accessibility levels.

The most conservative estimates for the difference between the school and racial gaps in Combinatorics and non-Combinatorics problems are -3.7% (Table 4.4, Model 4, significant at the 1% level) and -1.5% (Table 4.5, Model 4, significant at the 5% level), respectively. The regression models indicate, thus, that after controlling for question difficulty, the private-public school difference in the rate of correct responses in Combinatorics problems is 3.7 percentage points smaller than in non-Combinatorics problems - and that the black-white gap is 1.5 percentage points smaller in Combinatorics than in non-Combinatorics problems. These are sizable fractions of the original, unconditional gaps. The overall private-public raw gap is 12.9% across all problems in the sample. The final estimate of the gap within Combinatorics problems, after accounting for item difficulty levels, is thus 9.2% (12.9% - 3.7%), which is about a third smaller than the original raw gap. There is also a sizable reduction in the racial gap across Combinatorics problems. The overall raw racial gap is 5.9%, while across Combinatorics problems, it is estimated as 4.4% (5.9% - 1.5) - about a quarter smaller.

Table 4.4: Regressions of private-public school achievement gaps (an indicator of accessibility, in the opposite direction) on a Combinatorics dummy and covariates.

	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>	<i>Model 4</i>
<i>(Intercept)</i>	0.131*** (0.004)	0.083*** (0.013)	0.269*** (0.042)	0.033 (0.047)
<i>combinatorics</i>	-0.060*** (0.017)	-0.060*** (0.016)	-0.059*** (0.016)	-0.037** (0.013)
<i>item position</i>			-0.001*** (0.000)	
<i>difficulty</i>				0.723*** (0.147)
<i>difficulty sq.</i>				-0.861*** (0.155)
<i>year fixed effects</i>	<i>no</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
<i>R-squared</i>	0.017	0.105	0.136	0.416
<i>N (items)</i>	628	628	628	628

Significance: *** = $p < 0.001$; ** = $p < 0.01$; * = $p < 0.05$. Standard deviations are shown in parenthesis, below each coefficient.

Table 4.5: Regressions of black-white achievement gaps (an indicator of accessibility, in the opposite direction) on a Combinatorics dummy and covariates.

	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>	<i>Model 4</i>
<i>(Intercept)</i>	0.060*** (0.002)	0.038*** (0.007)	0.138*** (0.022)	0.087*** (0.021)
<i>combinatorics</i>	-0.027** (0.009)	-0.029*** (0.009)	-0.029*** (0.008)	-0.015* (0.006)
<i>item position</i>			-0.001**** (0.000)	
<i>difficulty</i>				0.205** (0.066)
<i>difficulty sq.</i>				-0.356*** (0.052)
<i>year FE</i>	<i>no</i>	<i>yes</i>	<i>Yes</i>	<i>yes</i>
<i>R-squared</i>	0.015	0.081	0.114	0.550
<i>N</i>	628	628	628	628

Significance: *** = $p < 0.001$; ** = $p < 0.01$; * = $p < 0.05$. Standard deviations are shown in parenthesis, below each coefficient.

Although it is challenging to assess the practical significance of the smaller gaps found across Combinatorics problems, prior studies can provide an illustration of the relationship between ENEM scores and university admissions. Based on internal data from the Federal University of Minas Gerais, a selective public university in Brazil, Winther and Golgher (2010) simulated how a hypothetical bonus-score policy in ENEM scores would impact the sociodemographic profile of admitted students. In their simulations, providing non-white students with 5% bonus scores in ENEM would increase the share of non-white admitted students at UFMG from 29.2% to 46.0% in the university's most selective courses. Their simulations also indicate that the share of admitted students coming from public high schools would increase from 21.9% to 26.3%. Based on these simulations, the results obtained in this dissertation study – in particular the 3.7% reduction in school gap across Combinatorics problems – seem to have practical significance for higher education admissions. That is, if *all* problems in the ENEM had school gaps as small as the gaps observed in Combinatorics problems, we should expect a more equitable pool of admitted students in highly selective universities in Brazil (i.e. increased share of students who graduated from public high schools and increased share of non-white students).

Checking for the robustness of the accessibility regressions results is especially important in light of the fact that the data shows that the confounding effect of item difficulty is relevant (i.e. the numerical estimates of accessibility varied once this variable was included as a covariate). As discussed in the previous chapter, methods that more explicitly take the confounding variable (difficulty) into account can contribute to checking for the consistency of the results. To do that, I rerun all models used to analyze accessibility using matching techniques. I used two variables for selecting the matched samples: item difficulty and item

position, as well as the imposition that each pair of items belongs to the same exam. This procedure generated a subsample of the original 628 ENEM problems, now comprised of 56 observations (the original 28 Combinatorics problems and 28 non-Combinatorics problems assigned as their matching pairs).

The matched samples model results are consistent with the original regressions that used the full database of ENEM problems (Tables C.1 and C.2 provided in Appendix C). Even with the smaller sample size, all coefficients of interest remain significant at the 1% or 5% levels in the most conservative model with full controls (Model 4). Compared to the previous models, there is little variation in the values of the estimated coefficients. The most conservative estimates using the matching technique indicate that Combinatorics problems, when compared to non-Combinatorics problems, present a 3.8% smaller school gap (this estimate was 3.7% in the non-matched regression) and a 2.1% smaller racial gap (this estimate was 1.5% in the non-matched regression). The overall conclusion that Combinatorics problems seem to be more accessible (smaller gaps) than non-Combinatorics items remains the same.

Research Question 2

Research Question 2 asks “What (if any) are some characteristics of Combinatorics problems that are associated with their difficulty and accessibility levels?”. To approach this question from a quantitative perspective, I re-run previous regression models including two quantifiable characteristics of the problems: question length (word count) and notation (a dummy variable indicating whether the problem statement includes complex mathematical notation).

As a starting point for the analysis, Table 4.6 shows descriptive statistics about the two variables for Combinatorics and non-Combinatorics problems.

Table 4.6: Quantifiable characteristics of Combinatorics and non-Combinatorics items used in the study.

	Combinatorics problems	Non-Combinatorics problems	All problems
Length (average word count)	86.2	91.6	91.4
Notation (% problems with mathematical notation)	35.7%	15.7%	16.6%
N	28	600	628

Combinatorics items of the ENEM exam have, in general, shorter problem statements than non-Combinatorics items (averages are 86.2 words and 91.6 words, respectively). In this comparison, thus, problem statements in Combinatorics seem to have a tendency to be “simpler” (Daroczy et al., 2015). On the other hand, and in contrast to descriptions in the literature discussed in the previous chapter, Combinatorics problems in ENEM include more complex mathematics notation than non-Combinatorics problems. While 35.7% of Combinatorics problems included some of the notation considered complex in the coding scheme used in this study, this kind of notation was only present in 15.7% of non-Combinatorics problems. Therefore, in regard to notation, Combinatorics problem statements in ENEM seem to be more complex than problems in other domains. This issue will be further explored in the qualitative portion of this study.

Resorting back to regressions, Table 4.7 reports the results of models that include the two quantified characteristics of problems to the previously estimated models of difficulty (Table 4.3, Model 3), school gap (Table 4.4, Model 4), and racial gap (Table 4.5, Model 4). To facilitate comparisons, the original estimates from the previous models are reproduced in Table 4.7 alongside the new results. For the sake of simplicity, year fixed effect coefficients are omitted from the table, even though they were included in all regressions.

**Table 4.7: New regression models including observable characteristics
(length and notation).**

	<i>Difficulty</i>		<i>School gap</i>		<i>Race gap</i>	
	<i>Previous</i>	<i>New</i>	<i>Previous</i>	<i>New</i>	<i>Previous</i>	<i>New</i>
<i>(Intercept)</i>	0.443*** (0.061)	0.397*** (0.061)	0.033 (0.047)	0.131* (0.052)	0.087*** (0.021)	0.122*** (0.024)
<i>combinatorics</i>	0.051* (0.023)	0.042 (0.023)	-0.037** (0.013)	-0.037** (0.013)	-0.015* (0.006)	-0.014* (0.006)
<i>item position</i>	0.002*** (0.000)	0.002*** (0.000)			0.087*** (0.021)	
<i>length</i>		0.000*** (0.000)		0.000*** (0.000)		0.000*** (0.000)
<i>notation</i>		0.063*** (0.013)		-0.017* (0.007)		-0.010** (0.003)
<i>difficulty</i>			0.723*** (0.147)	0.848*** (0.146)	0.205** (0.066)	0.245*** (0.066)
<i>difficulty sq.</i>			-0.861*** (0.155)	-0.930*** (0.114)	-0.356*** (0.052)	-0.375*** (0.052)
<i>year fixed effects</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
<i>R-squared</i>	<i>0.105</i>	<i>0.153</i>	<i>0.416</i>	<i>0.446</i>	<i>0.550</i>	<i>0.566</i>
<i>N (items)</i>	<i>628</i>	<i>628</i>	<i>628</i>	<i>628</i>	<i>628</i>	<i>628</i>

Significance: *** = $p < 0.001$; ** = $p < 0.01$; * = $p < 0.05$. Standard deviations are shown in parenthesis, below each coefficient. All regressions include year fixed effects, even though their coefficients are not reported.

I begin the analysis of the results from Table 4.7 by inspecting the direction and significance of the coefficients of each new variable. First, variable *length* has a statistically significant positive association with difficulty and with achievement gaps but has a numeric value close to zero in all regressions. Thus, although the estimates indicate that longer statements are associated with higher difficulty and larger gaps, this association seems to be of negligible practical importance.

The variable *notation*, on the other hand, has coefficients of arguably sizeable magnitude, which are all significant at the 5% level (or lower). This variable correlates positively with *difficulty*, indicating that the presence of notation is associated with 6.3% *higher* incorrect

response rates. This result is in line with the literature discussed in the previous chapter (Daroczy et al., 2015).

Notation correlates negatively, though, with school and racial gaps. The presence of notation is associated with a 1.7% *lower* school gap and a 1.0% lower racial gap. This result contradicts a hypothesis derived from the literature - as discussed in the previous chapter, the literature implies that problems with simpler statements are expected to be associated with increased accessibility. This puzzling result will be further investigated in the qualitative part of this study.

Table 4.7 can also be examined by observing how the inclusion of new variables alters the combinatorics dummy coefficients. This inspection can help illuminate whether the results found in answering Research Question 1 (i.e. ENEM Combinatorics problems present higher difficulty and smaller school and racial gaps) are attributable to the new variables (length and notation). For the case of the school gap proxy, the difference in gaps between Combinatorics and non-Combinatorics problems was 3.7% before the inclusion of the new variables, and 3.6% after. For the case of the racial gap proxy, the difference in gaps between Combinatorics and non-Combinatorics problems was 1.5% before the inclusion and 1.4% after. Thus, in regard to accessibility, results from Table 4.7 indicate that the inclusion of the item-specific characteristics does not change the estimated coefficients for the combinatorics dummy. In other words, these characteristics - as captured by the simple indicators used in this study - seem not to be the important factors explaining the differential levels in gaps documented between Combinatorics and non-Combinatorics problems in the ENEM data.

Regarding difficulty levels, there is a more pronounced change in the combinatorics dummy coefficient (from 5.1% to 4.2%), and it ceases to be significant. The results indicate,

thus, that the presence of complex notation is a candidate explanatory factor for the difference in difficulty levels observed between Combinatorics and non-Combinatorics problems.

It is also worthwhile to inspect whether the two linguistic factors investigated—item length and item notation—exhibit different explanatory patterns between Combinatorics and non-Combinatorics problems. To accomplish this, I independently replicate the results of Table 4.7 for the subsamples of both Combinatorics and non-Combinatorics problems. The results from replicating the regression with *length* and *notation* are provided for Combinatorics (Table C.4) and non-Combinatorics problems (Table C.3) are provided in Appendix C. For the subpopulation of non-Combinatorics problems, the results are in line with the results obtained for the whole sample of problems. All coefficients of interest have the same direction as in Table 4.7 and are significant at the 5% level. For the subsample of Combinatorics problems, none of the coefficients in the regression are significant, likely due to small sample sizes.¹¹

Research Question 3

This dissertation's third and last research question inquires about how students interact with combinatorics problems. Research Question 3 is stated as follows:

(3) What are some of the ways in which students access problems in Combinatorics? What are some of the ways in which students fail to provide correct answers to problems in Combinatorics? What is the relationship between difficulty and accessibility in students' interactions with Combinatorics problems?

I approach this research question qualitatively, using data from the interviews I conducted with students from a preparatory school for the ENEM exam. In comparison to the

¹¹ In this case, the size of the sample becomes much smaller ($n=28$ combinatorics problems) than in previous estimations ($n=628$ for the full sample and $n=600$ for the non-Combinatorics sample). Accordingly, standard errors of estimated coefficients become much larger than in previous estimates.

quantitative analysis in the previous section, the qualitative analysis presented here intends to be both more explanatory and more exploratory. This investigation is intended as explanatory in the sense that beyond looking for quantitative patterns in data, I aim to unveil insightful examples of specific ways in which students interact with Combinatorics and non-Combinatorics problems. The analysis is exploratory in the sense that I aimed to explore the richness of information available in the interviews by being attentive to unpredicted phenomena related to the hypotheses raised in Chapter 2.

The results presented in this section are based on a small set of interviews with a particular group of students and a restricted number of problems and should thus be interpreted with parsimony. My goal is not to provide any definitive, generalizable answers to the research question proposed but instead to provide initial insights into potentially relevant phenomena. I interpret the episodes presented in this section as “proofs of existence” of *some ways* in which students interact with Combinatorics problems.

To provide an overview of the outcomes of the interviews, Table 4.8 summarizes the data on the first round of coding that comprised indicators of accessibility and difficulty across Combinatorics and non-Combinatorics problems.

Table 4.8. Indicators of accessibility and difficulty in student interviews.

<i>Accessibility</i>								
	<i>Self-report (I understand what this problem asks)</i>		<i>Researcher rating (Student rephrasal of what the problem asks)</i>		<i>Engagement with mathematics</i>			<i>n</i>
	<i>yes</i>	<i>partial/no</i>	<i>yes</i>	<i>partial/no</i>	<i>none</i>	<i>procedural</i>	<i>conceptual</i>	
<i>Combinatorics</i>	88%	13%	70%	30%	19%	20%	61%	96
<i>non-Combinatorics</i>	92%	8%	54%	46%	8%	29%	63%	48
<i>Overall</i>	89%	11%	65%	35%	15%	23%	62%	144

<i>Difficulty</i>					
	<i>Correct answer choice</i>		<i>Valid mathematical justification for correct answer choice</i>		<i>n</i>
	<i>yes</i>	<i>no</i>	<i>yes</i>	<i>no</i>	
<i>Combinatorics</i>	41%	59%	22%	78%	96
<i>non-Combinatorics</i>	27%	73%	25%	75%	48
<i>Overall</i>	36%	64%	23%	77%	144

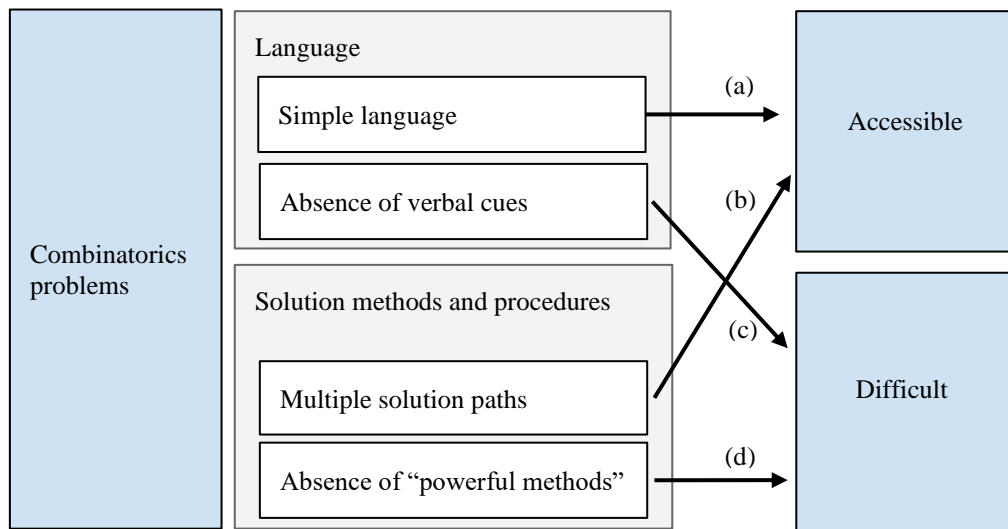
Results from Table 4.8 indicate that despite understanding what most problems asked, students encountered challenges in solving the proposed problems. While, on average, students rephrased problems in valid ways 65% of the time and engaged in conceptual problem-solving attempts 62% of the time, only 36% of their answer choices were correct. Overall, these results anticipate that accessibility and difficulty coexisted in students' interactions with the proposed problems.

Table 4.8 also indicates some initial differences between how students interacted with Combinatorics and non-Combinatorics problems. Combinatorics problems were correctly rephrased at higher rates than non-Combinatorics problems (70% and 54%, respectively). Thus, considering the first part of the definition used in this study for *accessibility* (i.e. understanding what problems ask), Combinatorics problems in the interviews were more accessible than non-

Combinatorics problems. Nevertheless, the higher rates of student understanding across Combinatorics problems were not translated into higher rates of conceptual engagement with mathematics in students' solution attempts to solve the problems. On average, students engaged in conceptual reasoning in 61% of their interactions with Combinatorics problems and in 63% of their interactions with non-Combinatorics problems.

Also, when I asked students to explain the reasoning behind their answers, I classified the explanations as valid 22% of the time across Combinatorics problems and 25% across non-combinatorics problems. The low rate of correct responses (36%) and explanations (23%) across all problems in the interviews indicates that, in general, students found the problem set very challenging. We can also infer from Table 4.8 that some of the correct answers were not backed up by student understanding of the mathematics involved in the problem, an issue that will be discussed further in this section. Nevertheless, students were very collaborative in making an effort to explain their solution attempts, even when those attempts were incomplete.

The resulting themes from the second round of analysis, in which I looked for phenomena that help exemplify how students interact with Combinatorics and non-Combinatorics problems, are summarized in Figure 4.3. This analysis was initially inspired by hypotheses raised in Chapter 2, and the final themes are thus presented in association with them. In the remainder of this section, I explore each theme in the sequence presented in Figure 4.3 and then revisit Research Question 3 to discuss how the qualitative findings help shed light on them.



<i>Connection</i>	<i>Theme</i>
<i>a) Relationship between the simplicity of language and accessibility</i>	<ul style="list-style-type: none"> • <i>Less technical vocabulary and symbols allowing students to understand what problems ask.</i>
<i>b) Multiple solution methods and accessibility</i>	<ul style="list-style-type: none"> • <i>Use of context as leverage for understanding.</i>
<i>c) Relationship between the absence of verbal cues and difficulty</i>	<ul style="list-style-type: none"> • <i>Trying to use methods in situations in which they are not appropriate</i>
<i>d) Relationship between the absence of powerful methods and difficulty</i>	<ul style="list-style-type: none"> • <i>Strategic guessing</i>

Figure 4.3. Coding themes used in the qualitative analysis.

a) Relationship between the simplicity of language and accessibility: less technical vocabulary and symbols allowing students to understand what problems ask

I start this subsection with an overview of the symbols and vocabulary that I identified as potentially challenging in the interview problem set. This list was built prior to the interviews in consultation with two experienced high school teachers in Brazil and is reproduced in Table 4.9. We read all 12 problems together and searched for words and symbols that (i) may not be known to students without a strong mathematical background and (ii) were important to understand and solve the problem. This procedure is more nuanced than the procedure used in quantifying the

presence of mathematical notation and symbols in Research Question 2. First, it is more comprehensive, as it allows for words (as well as symbols) to be considered challenging for students. The prior procedure used in Research Question 2 took only symbols (not words) into account. Second, in this qualitative analysis, I kept track of these terms and symbols separately in regard to whether they were present in the problem statement or the problem answer choices. In the quantitative analysis, in contrast, technical symbols that appeared either in the problem statement or in the problem's answer choices were indistinguishably accounted for (i.e. a problem was coded as containing potentially challenging mathematical notation if a potentially challenging symbol was present in the question statement and/or in the answer choices).

Table 4.9. Pre-identified potentially challenging words and symbols.

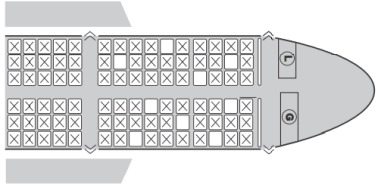
	<i>Content area</i>	<i>Pre-identified mathematical vocabulary and symbols in question statement</i>	<i>Student rephrasing (% correct)</i>
<i>Pair of problems A</i>			
<i>Problem 1- Detective game</i>	<i>Combinatorics</i>	-	42%
<i>Problem 2- Container volume</i>	<i>Geometry</i>	-	67%
<i>Pair of problems B</i>			
<i>Problem 3 - Bank password</i>	<i>Combinatorics</i>	<i>Ratio</i>	67%
<i>Problem 5- Teacher hiring</i>	<i>Data analysis</i>	<i>Average</i>	67%
<i>Pair of problems C</i>			
<i>Problem 7- Father Drawing</i>	<i>Combinatorics</i>	-	92%
<i>Problem 8- Scouts' tent</i>	<i>Geometry</i>	<i>Right prism, orthogonal projection, plane</i>	25%
<i>Pair of problems D</i>			
<i>Problem 12 - Paths on map</i>	<i>Combinatorics</i>	-	83%
<i>Problem 11- Motorbike trip</i>	<i>Algebra</i>	-	58%
<i>Other Combinatorics problems</i>			
<i>Problem 4 - Renting movies</i>	<i>Combinatorics</i>	-	50%
<i>Problem 6 - Aitplane seats</i>	<i>Combinatorics</i>	-	100%
<i>Problem 9 - Volleyball teams</i>	<i>Combinatorics</i>	-	92%
<i>Problem 10- Toy train</i>	<i>Combinatorics</i>	<i>Combinations</i>	33%

Overall, few interview problems contained complex words or symbols. I identified potentially challenging linguistic markers in only four out of the twelve problem statements. Two of these were Combinatorics problems, one was a Geometry problem, and one was a Data Analysis problem.

I present evidence about the relationship between the use of simple language statements and access by contrasting two Combinatorics problems in which students presented varying levels of correctness in their explanations of what the problems asked them to do. This evidence is intended to illustrate and expand a core hypothesis drawn from prior literature on Combinatorics discussed in Chapter 2 - descriptions of Combinatorics problems being accessible frequently point that problems in this domain make less use of complex mathematical concepts or notation (e.g., Lockwood et al., 2020; Sandefur et al., 2022).

Problem 10, Toy Train, was the problem with the lowest rate of adequate restatement by students. I considered that only three students explained it consistently. On the other extreme, in Problem 6, Airplane Seats, I considered that the problem was correctly rephrased by all twelve students. Problem 6 is reproduced below.

A family of seven adults, after deciding the itinerary of their trip, consulted the website of an airline and found that the flight for the chosen date was almost full. In the figure available on the website, the seats occupied are marked with an X, and the only seats available are those shown in white.



The number of different ways to accommodate the family on that flight is given by

a) $\frac{9!}{2!}$ b) $\frac{9!}{7! \times 2!}$ c) $7!$ d) $\frac{5!}{2!} \times 4!$ e) $\frac{5!}{4!} \times \frac{4!}{3!}$

Figure 4.4. Problem 6, Airplane Seats.

Note that even though the symbol for factorials appears in the problem's answer choices, no technical symbol or vocabulary is present in Problem 6's statement. Students could rephrase it in their own words in arguably valid ways, demonstrating they understood what the problem asked. The following excerpts exemplify how some of the students explained what the problem asked.¹²

Student 3: There are 7 people and 9 seats. He wants to know the possibilities of... how many ways can the people sit.

Student 5: He is asking me to know in how many ways can I accommodate the people in this family. [In how many ways] can I put 7 people in this plane, which has only 9 available seats?

Student 6: A family will have an airplane trip. There are 9 vacant seats and they are 7 people. It's asking me in how many ways this family can be accommodated in this airplane with these 9 empty seats.

Student 10: There are 7 people and there is an airplane... an airplane, right? Then there are some seats spread out [on the plane]. In the question, he wants to know what is the formula... No, not the formula... The calculation that needs to be done to figure out how many different ways the people can accommodate themselves.

I considered that all students were able to explain what the problem asked, even if with varied degrees of preciseness. Student 6, for example, provides a complete explanation of the problem, including all relevant numeric information. Student 10, on the other hand, provides a

¹² Original quotes:

Student 3: *São 7 pessoas aí são 9 assentos. Ele quer saber a possibilidade de, tipo, de quantas formas podem sentar as pessoas?*

Student 5: *Ele está me pedindo para saber de quantas formas eu posso estar acomodando as pessoas dessa família... que eu posso por 7 pessoas nesse vôo que só tem 9 assentos disponíveis.*

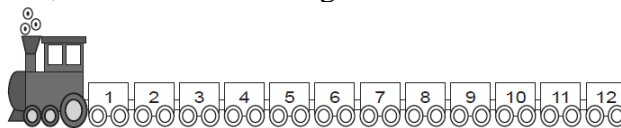
Student 6: *Uma família vai fazer uma viagem de avião. Tem 9 lugares vagos e eles são 7 pessoas. Está perguntando de quantas formas essa família pode se acomodar nesse avião com esses 9 lugares vazios.*

Student 10: *É, tem 7 pessoas aí tem um avião... um avião, não é? Aí tem algumas poltronas espalhadas. Na pergunta ele quer saber qual é a fórmula, ... É, fórmula, não. O cálculo que faz para descobrir quantas formas diferentes dá pra cada pessoa se acomodar.*

less complete explanation that omits the number of available seats. Even so, the student manages to explain that there are a given number of family members and some available seats, and that the question is asking how many different ways exist to arrange these people in the available seats. As per the criteria used to code students' descriptions of problems (presented in Chapter 3), enough information is provided for a listener with knowledge of Combinatorics to understand the mathematical situation proposed in the problem.

Problem 10, Toy Train, is on the opposite side of the spectrum in terms of student understanding of what the problem asked, with only four students correctly paraphrasing it, as per my coding. The problem is reproduced below.

A company makes toy trains comprised of a black locomotive and 12 wagons numbered 1 to 12. Of the 12 wagons, 4 are painted red, 3 are painted blue, 3 are painted green, and 2 are painted yellow. Trains are assembled starting with the locomotive and attaching 12 numbered wagons in increasing order, as shown in the image below.



The number of different trains that can be made, as expressed by combinations, is given by

- A** $C_{12}^4 \times C_{12}^3 \times C_{12}^3 \times C_{12}^2$
- B** $C_{12}^4 + C_8^3 + C_5^3 + C_2^2$
- C** $C_{12}^4 \times 2 \times C_8^3 \times C_5^2$
- D** $C_{12}^4 + 2 \times C_{12}^3 + C_{12}^2$
- E** $C_{12}^4 \times C_8^3 \times C_5^3 \times C_2^2$

Figure 4.5. Problem 10, Toy Train.

The following excerpts illustrate what I considered to be insufficient explanations of what the problem asks.¹³

¹³ Original quotes:

Student 5: It's asking the variations... of where could the trains be painted... No, no, sorry. It's not the color, it's the quantity and how many trains can be assembled. He wanted to know what is... what is the combination.

Student 11: It's asking for the calculation that would be the possible variations... But, [I] couldn't arrive at the calculation. [Researcher: But can you explain to me what the problem is asking without doing the calculation?] I think that it's asking to indicate how many possibilities of colors each wagon can be.

I considered the two explanations in the examples to be insufficient because they do not convey enough information for the listener to recreate the mathematical situation originally provided in the problem. For instance, while Student 5 points out that the problem deals with variations of the train colors, their problem restatement doesn't describe in full how this variation occurs. Also, Student 6's explanation of what the problem asks focuses on the colors of *each wagon* rather than the different *train designs* - as originally asked in the question.

For this problem, the two consulted high school teachers and I anticipated that the term "combinations" could be a challenging term in the problem statement. Because of that, I asked each student to explain to me what this word meant. Note that the meaning of "combinations" in this problem is indeed precise and mathematics-specific. It means combinations in its "combinatorial" sense - among n different objects, how many different sets of k elements can be constructed. In the interviews, many students revealed a lack of knowledge about the mathematics-specific meaning of the word "combinations". When I asked students what this

Student 5: Desse exercício eu não consegui entender muito. Mas ele está pedindo as variações de onde poderia ser pintados os trens? Não, não perdão. Não é coloração dos trens, é a quantidade de quantos trens poderiam ser montados. E ele queria saber qual é a... qual é a combinação?

Student 11: Está pedindo para ver a conta que que seria das possíveis variações, mas, não teria como chegar a conta, não. [Researcher: Mas consegue me explicar o que está pedindo pelo menos, sem fazer a conta?] Eu acho, né, que está pedindo para falar quantas possibilidades de cor pode ser o cada vagão.

term meant, eight students provided definitions that I considered to be incomplete. The following excerpts exemplify this issue.¹⁴

Student 5: *It's the number of ways in which... it's how many ways it can be done...*

Student 7: *Hum, I'd say combination would be this that I commented about... How many ways you can build something.*

Student 10: *It's how many ways you can do. How many, how many ways you can do each train.*

Student 11: *I think it's the different ways.*

Even though all four students in the passage tried to connect the term “combinations” to some idea in the realm of counting (e.g., “the number of ways”, “how many ways”), only partial understandings of the meaning of this concept can be inferred from their explanations. None of the students in the passages mentions, for example, how the “combinations” are to be made (e.g. ordered, unordered, with or without repetition). For these students, thus, it seems likely that a lack of prior knowledge of the mathematical concept of combinations may have acted as a barrier to understanding what the problem asked.

The above analysis provided examples that illustrate the argument made in the literature about the connection between language and accessibility *within* Combinatorics problems. The examples illustrate the idea that Combinatorics problem statements with less technical vocabulary and notation tend to be more accessible. That is, the less mathematics-specific vocabulary is used in the problem statement, the higher the chance that students will understand

¹⁴ Original quotes:

Student 5: *É de quantas formas pode ser... de quantas formas pode ser feita...*

Student 7: *Ah, eu diria que a combinação seria isso que eu comentei... De quantas formas pode ser montado algo.*

Student 10: *É de quantas formas é para fazer. De quantas, de quantas formas da para fazer cada trem.*

Student 11: *Acho que as formas diferentes.*

what the problem asks. While this argument is frequently made to explain why Combinatorics problems may be more accessible than problems in other domains, the examples provided here highlight that this factor may also be important in determining the accessibility levels *within* Combinatorics problems.

b) Multiple solution methods and accessibility: use of context as leverage for understanding

The literature discussed in Chapter 2 indicates that a possible explanatory factor for Combinatorics problems to be particularly accessible is that they allow for multiple solution methods (e.g., English, 2005; Hurdle et al., 2016). Although there were some cases in the interviews in which students engaged in varied strategies to try to solve some of the problems, these instances were not particularly common. Because of that, I broaden the discussion on how Combinatorics problems may allow for multiple entry points with the introduction of a related phenomenon identified in the data. In some combinatorics problems, I identified that students seemed to leverage the context of the problem to understand what was being asked.

There is thus a slight shift in the analysis provided in this section in comparison to the hypotheses drawn from the literature. English (2005) and Hurdle et al. (2016) highlighted that the availability of multiple solution paths may be linked to student engagement with valuable mathematical processes when trying to solve problems in Combinatorics. Here, though, I stress how multiple approaches to a problem can be related specifically to the “initial” phase of accessibility - understanding what a problem asks. That is, instead of looking into multiple ways of solving problems, I highlight that Combinatorics problems, in particular, seem to allow students to use context to understand what was being asked. In principle, this may seem a truism - all problems in the interviews have contexts, so students will have to talk about these contexts

to explain what problems ask. The details about how students did this in Combinatorics problems, though, reveal interesting patterns in the particularities of this phenomenon.

First, it appeared that students were able to apprehend all pertinent information from the Combinatorics problems. That is, when they explained the problems in their own words, it was less likely that important information would be omitted. In some Combinatorics problems, this might be related to the problem statements being particularly short. That is the case of Problem 6, Airplane Seats, discussed above, and Problem 7, Painting Circles. An additional example from the Airplane Seats problem exemplifies how students successfully explained what the problem asked, using its context.¹⁵

Student 12: I understood there are nine available seats for seven people. But then people can kind of swap places. So, there are different ways they can accommodate themselves. In how many different ways can they seat?

In Student 12's response, they are able to recall all parameters of the problem using their contextual meanings (seven people, nine seats). The Airplane Seats and the Painting Circles problems are among the three shortest problems in the problem set with 70 and 60 words, respectively, compared to an average of 104 words across the 12 interview questions. The combination of a relatable context and a short problem length may thus have helped students make sense of what the problems asked and be able to recall all relevant information when trying to rephrase them.

More surprisingly, students could also summarize the relevant information in compact ways in Combinatorics problems with longer statements by using relevant contextual

¹⁵ Original quotes:

Student 1: *Tá falando de um avião em que só tem 9 poltronas vazias e precisa colocar uma família de 7 pessoas nessas poltronas vazias. E a pergunta é de quantas formas essa família pode se acomodar no voo.*

Student 12: *Entendi que são 9 lugares disponíveis para 7 pessoas. Só que aí essas pessoas podem, tipo, revezar de lugar. Então tem diferentes formas delas se acomodarem. [...] Quantas formas diferentes eles podem se sentar?*

information appropriately. A passage about Problem 1, Detective Game (reproduced below) exemplifies this phenomenon.¹⁶

A school principal invited 280 third-year students to participate in a game. Suppose there are 5 objects and 6 characters in a house with 9 rooms; one of the characters hides one of the objects in one of the rooms of the house. The game's objective is to guess which object was hidden by which character and in which room of the house. All students decided to participate. Each time, a student is drawn randomly to try to answer the challenge. The answers must always be different from all of the previous ones, and the same student cannot be drawn more than once. If the student's answer is correct, the student is declared the winner, and the game ends. The principal knows that a student will eventually provide the correct answer because there are

- A) 10 more students than possible different answers.
- B) 20 more students than possible different answers.
- C) 119 more students than possible different answers.
- D) 260 more students than possible different answers.
- E) 270 more students than possible different answers.

Figure 4.6: Problem 1, Detective Game.

Student 8: There are 5 objects, 6 people, and a house with 9 rooms. Then, one of the persons hides an object. This is the game. And then the 280 students will participate. But the answers always need to be different, and each one [of the students] can only be randomly picked once.

The problem has the longest statement in the problem set with 119 words. Even so, Student 8 could describe the problem briefly in their own words, including all the relevant numeric information. In contrast to this example of concise and complete Combinatorics problem rephrasing, a different situation occurs when the same student works on Problem 11, Motorcycle Trip, a problem in Algebra. The problem involves proportions and is reproduced below.

¹⁶ Original quote: *São 5 objetos, 6 personagens e uma casa com 9 cômodos. Aí um dos personagens esconde um objeto, essa é a brincadeira. E aí os 280 alunos vão participar. Só que a resposta sempre tem que ser diferente e cada um só pode ser sorteado de uma vez.*

A motorcyclist plans to take a trip with a destination 500 km from his home. His motorcycle consumes 5 liters of gasoline for every 100 km traveled, and the motorcycle's tank has a capacity of 22 liters. On the map, he noted the last available refueling station on the journey, called Estrela, is 80 km from his destination. He intends to leave with a full motorcycle tank and plans to make only two stops for refueling, one on the way out and the other on the way back, both at the Estrela gas station. When refueling for the outward journey, he should also consider enough fuel to ride 200 km in the destination city. The minimum amount of fuel, in liters, that this motorcyclist must refill at the Estrela gas station on the outward journey, which is sufficient to make the second refueling, is

- a) 13. b) 14. c) 17. d) 18. e) 21.

Figure 4.7: Problem 11, Motorbike Trip.

This problem is also long, with 106 words (the fourth longest in the set). When trying to restate the problem in their own words, Student 8 revealed that they found it challenging to keep track of all relevant information.¹⁷

Student 8: So, if I were doing an exam, I would highlight the most important information so that I don't get lost. But the question he wants to know is the minimum amount of fuel. But in liters, right? Let me see if it's already shown in liters... yes, it's already in liters up there. That "the guy needs to refuel", he wants to know that.

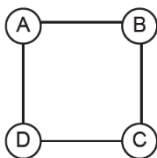
In the passage, the student spontaneously expresses a perception that keeping track of the information is challenging in this problem. The student proceeds to try to organize their answer and is ultimately unable to solve it after working on it for over 10 minutes. Two other students also spontaneously communicated a perception that this problem had a large amount of information that needed to be accounted for.

¹⁷ Original quote: *Então, se eu tivesse uma prova, eu ia a circular as informações mais importantes para mim não ficar perdida. Mas a questão é que ele quer saber a quantidade mínima de combustível - Só que ele quer em litros, né? Deixa eu ver se já está em litros... é, já está em litros lá em cima. Que o cara tem que reabastecer, ele quer saber isso.*

In sum, students seem to have been able to use the context of Combinatorics problems to understand and explain them in more compact and coherent ways than they did in the case of non-Combinatorics problems.

Another way in which students used the context to understand what was being asked relates to making sense of the underlying mathematics of problems. This phenomenon is exemplified in the excerpt below, in which Student 6 explains what is asked in Problem 7, Painting Circles.¹⁸

To stimulate his daughter's reasoning, a father made the following drawing and handed it to the child along with three pencils of different colors. He wants the child to paint only the circles and asks that those connected by a segment must have different colors.



In how many different ways can the child do what her father asks?

a) 6 b) 12 c) 18 d) 24 e) 7

Figure 4.8. Problem 7, Painting Circles.

Student 6: So, the father made a drawing for his daughter with four balls. The girl has three different colors, and she needs to paint four balls in such a way that the balls that are connected by... by the same segment, there, don't have the same color. So, for example, A cannot have the same color as B, neither D

In this passage, the student uses the context of the problem to provide an example that illustrates mathematically meaning information from the problem statement - the rule defining how the colors can be combined. The problem statement indicates that the daughter should paint

¹⁸ Original quote: *Então, um pai fez um desenho para a filha com 4 bolinhas. E a menina tinha 3 cores diferentes e ela tem que pintar as 4 bolinhas de modo que as que estão ligadas pelo mesmo... pelo mesmo segmento, ali, não tenham a mesma cor. Então, por exemplo, A não pode ter a mesma cor, que B, nem que a D.*

circles that are “connected by a segment” with the same color. The student used the drawing to correctly provide an example of what this mathematical constraint means¹⁹.

A different outcome occurs to the same student when working on Problem 8, Scouts' Tent, a similarly difficult problem (Problem 6 and Problem 8 comprise a pair of Combinatorics and non-Combinatorics problems of similar difficulty in the ENEM data), but in Geometry instead of Combinatorics. In this case, the student expresses confusion about the mathematical meaning of a concept (shortest path between points) provided in the problem statement. The problem is reproduced below, along with Student 6's explanation of what the problem asks.²⁰

A group of Boy Scouts was doing an activity in the park in the city where they live and set up a tent as shown in the photo of Figure 1. Figure 2 shows a schematic representation of the frame of this tent, which has the shape of a straight prism made of metal bars.

After the frame was assembled, one of the scouts observed an insect moving over them from vertex A towards vertex B, then from there towards vertex E, when finally it made the trajectory from vertex E to C. Assume that all these displacements were made the shortest path between the points. The projection of the insect's path on the plane that contains base ABCD is given by

Figure 4.9: Problem 8, Scouts' Tent.

¹⁹ The previously discussed passage in which Student 12 restates Problem 8, Airplane Seats, also contains an example of this phenomenon. The student uses the airplane seats situation to clarify the meaning of *different outcomes* in the problem's context by indicating that family members can “swap places”.

²⁰ Original quote: *Fui acompanhando o desenho. Então, tem a historinha dos escoteiros que montam a barraca. E aí tem as 2 imagens, a imagem da barraca, realmente, e a outra meio que do desenho ou do do formato dela. E aí está falando que passa e um inseto e faz o trajeto. Está perguntando qual das alternativas que representa esse trajeto feito por por inseto. Do caminho de menor distância entre os pontos. [Essa você sente que entendeu ela bem, mais ou menos, ou não?] Sim, entendi. Tá, só, esse "caminho de menor distância" fica meio, tipo, parece que foi colocado só para confundir porque, não tem meio que é outra distância menor do que aqui já é ali. Então, se ele não colocasse isso, também não acho que não faria muita diferença. Não entendi.*

Student 6: I was following the drawing. So there is the story of the Boy Scouts who assembled the tent. And there are the two images - the real image of the tent and the other that is kind of a drawing, or its shape. And then it says that an insect passes and makes the trajectory. It's asking which of the alternatives represents this trajectory made by the insect. About the shortest distance path between the points. [Interviewer: Do you think you understood this one, or you understood it partially, or didn't understand it?] Yes, I got it. It's just this "shortest distance path" that kind of looks like it was put there just to confuse, because there isn't kind of another distance rather than "here and there". Ok, if he [the problem author] had not put this in there I think it wouldn't make much of a difference. I didn't understand it.

In this passage, the student comments on what the problem's author meant by "minimum distance", and the context provided seems unhelpful in sorting out the confusion. Also, through the student's answer choice, evidence can be drawn that the student was unable to figure out the meaning of the concept of *orthogonal projection*, which is essential to understanding what the problem asks. The student chose an incorrect answer choice that depicted the path of the insect representing the "real word picture" perspective (answer choice B), instead of the path in the orthogonal projection that was asked in the problem statement (answer choice C).

The cases presented in this section highlight an expansive interpretation of the hypothesis that Combinatorics problems tend to be accessible because they allow for multiple solution paths. In the interviews, students used the context of Combinatorics problems as leverage to *understand what problems ask*. First, by using the problems' contexts, they could recall all the information provided and present it in compact ways. Second, they could bring meaning to mathematical concepts through context.

The results in this subsection complement the hypothesis drawn from the literature about the connection between the availability of multiple solution methods and accessibility in Combinatorics. Based on the cases discussed above, one can hypothesize that, in addition to

allowing for multiple solution methods, Combinatorics allows for multiple pathways - including the use of context - for students to make sense of what problems ask.

Before moving into themes related to problem difficulty (next two subsections), I briefly address an unpredicted empirical regularity regarding the “second part” of accessibility - student engagement levels with conceptual mathematics - which has not been discussed so far. As shown in Table 4.8, student engagement with conceptual mathematics was lower in Combinatorics problems than in problems from other domains. Overall, then, the higher level of student understanding of what problems ask in Combinatorics problems (70% in Combinatorics and 54% in non-Combinatorics problems) did not translate into higher levels of engagement with conceptual mathematics (61% in Combinatorics and 63% in non-Combinatorics problems). To shed light on this matter, the remainder of this subsection investigates three Combinatorics problems that presented high overall levels of accessibility. Looking closely at these problems can illuminate reasons why in their cases, specifically, higher student understanding *did* translate into higher engagement with conceptual mathematics.

In two such problems, students frequently used drawings in their annotations. This was the case of Problem 7, Painting Circles, in which I considered that 92% of students understood what the problem asked and 67% of the students engaged with conceptual mathematics. Student notes provided below exemplify how they used drawings to support their problem-solving attempts.

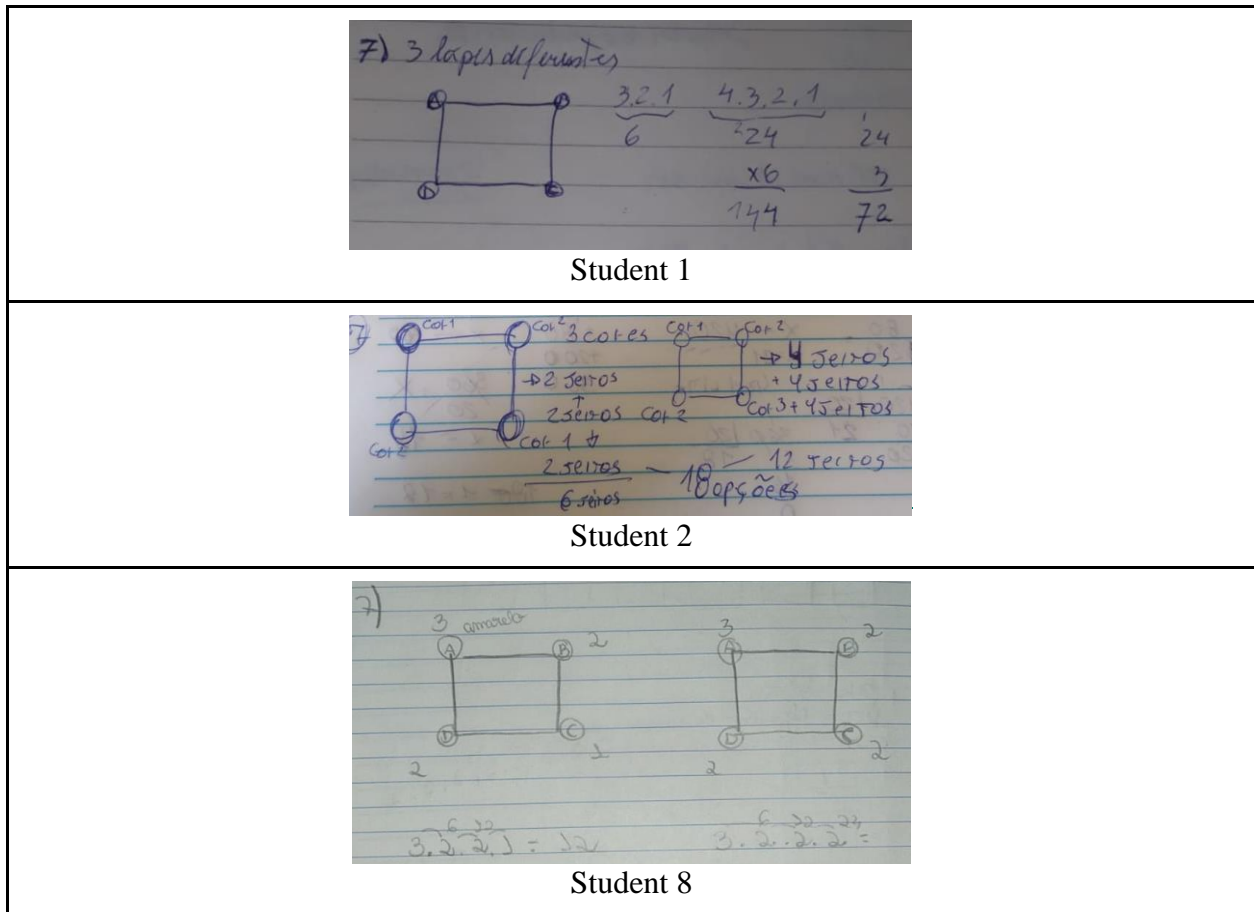


Figure 4.10: Students' annotations in Problem 7, Painting Circles.

The three students in the provided examples attempted to solve the problem with a sequential counting strategy, starting with circle A. In the work of all students, there is evidence of conceptual reasoning involving the multiplication principle. Their individual annotations indicate that once a circle is painted, there are fewer options available to paint the next circle (multiplying 3 x 2 and so forth). While among them only Student 2 ended up providing the correct answer choice, all students provided evidence of conceptual engagement with mathematics. Another contributing evidence that Problem 7 provided broad access to conceptual thinking is the fact that Student 1 (first image in Figure 4,10) was the student who had the lowest overall conceptual engagement across problems (33%), and nonetheless was able to demonstrate conceptual engagement in this problem. In Problem 7, thus, I interpret that students may have

benefited from an entry point that combined a visual representation and a relatively simple combinatorics concept (the multiplication principle).

Another problem in which students combined visual representations and a relatively simple combinatorics concept was Problem 12, Paths on Map, in which 75% of students showed evidence of conceptual engagement. The problem is reproduced below.

Three friends, André, Bernardo, and Carlos, live in a closed condominium in a city. The grid represents the location of the parallel and perpendicular streets, delimiting blocks of the same size in this condominium. The houses of André, Bernardo, and Carlos, respectively, are located at points A, B, and C.

André wants to go from his house to Bernardo's house without going through Carlos' house, following along the streets of the condominium, always moving to the right (\rightarrow) or upwards (\uparrow), according to the diagram in the figure. The number of different paths that André can use to carry out the displacement under the proposed conditions is

a) 4 b) 14 c) 17 d) 35 e) 48

Figure 4.11: Problem 12, Paths on Map.

Five students engaged with this problem by either drawing some valid paths that were to be counted or by visually reproducing them with their hands. They would continue to find new cases while trying to keep track of the number of different ways they had encountered. The combinatorial process involved in this strategy is thus also fairly simple, consisting of listing all distinguished cases and counting them out. Examples of student annotations and quotes are reproduced below.²¹

²¹ Original quotes:

Student 5: É, do A você segue reto [shows fingers pointing up]. Isso, e à direita. Seria, um caminho. Ou do A você vai para a direita, e depois sobe.

Student 8: Eu vou fazer isso olhando mesmo, eu vou fazer contando [...]. Eu consegui contar, acho que 14 mesmo.

Student 10: Aí eu vi eu enxerguei 4 possibilidades.

Student 12: Porque ele pode ir direto pela direita e subir - é um caminho. Aí ele pode subir um quadrante, virar à direita e subir de novo - são 2 caminhos. Ou ele pode subir até A, subir A inteiro e, na última linha, ele vira à direita.

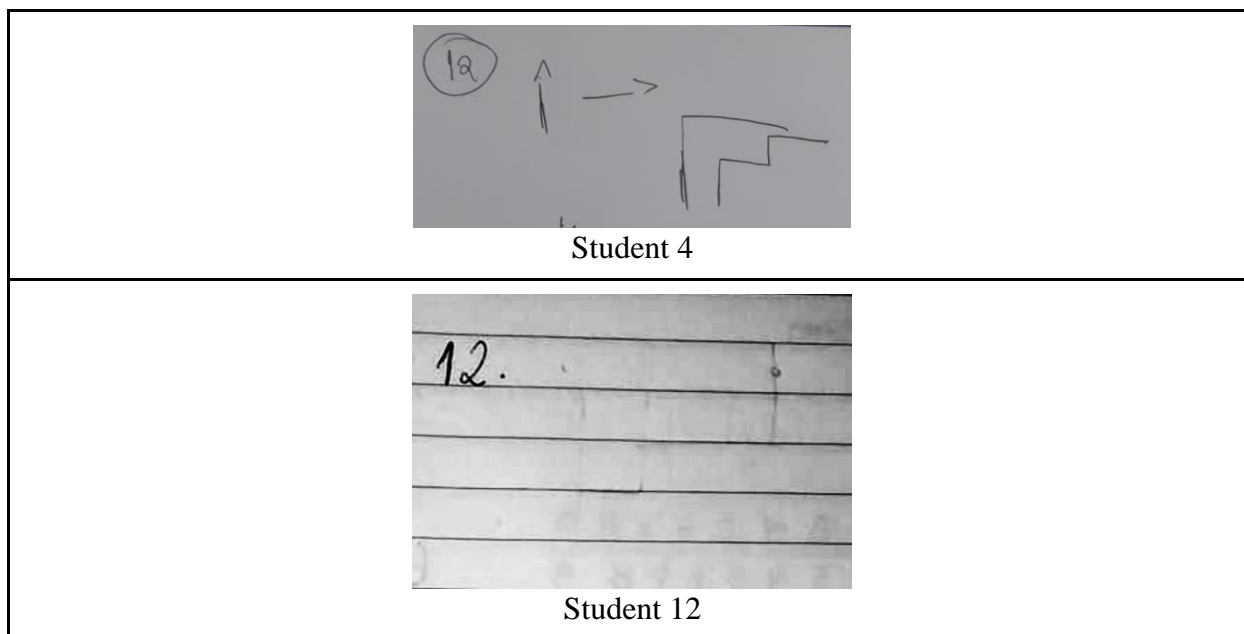


Figure 4.12: Students' annotations and quotes in Problem 12, Paths on Map.

Student 5: Well, from A you go straight [shows fingers pointing up]. That's it, and to right. This would be one path. Or from A, you turn right, and then go up.

Student 8: I'm going to do this by looking, I'm going to count [...]. I managed to count, I think it's 14 indeed.

Student 10: I looked and I saw 4 possibilities.

Student 12: Because he can go straight on the right and then go up - that's one path. Then he can go up one quadrant, turn right, and go up again - those are 2 paths. Or he can go up to A, go all the way up through A, and on the last line, he turns to the right.

Similarly to Problem 7, Painting Circles, evidence also indicates that the combination of a visual representation with a relatively simple combinatorial concept may have helped students in engaging conceptually with mathematics in Problem 12, Paths on Map. All students in the passages used a simple counting process, combined with visual representations in varied ways. Students 4 and 12 actually drew some of the possible paths, Student 5 used hand gestures, and Student 8 mentally visualized possible paths.

Similarly to the previous two problems, in Problem 3, Bank Password, students also frequently engaged with a simple mathematics concept, even though in this case without resorting to visual representations. Compared to the previous two problems, Problem 12 actually allowed students to engage with a simpler mathematical concept - sums. Problem 3 was the Combinatorics problem with the highest conceptual engagement levels (75%) and is reproduced below.

A bank requires its customers to create a six-digit password using numbers 0 to 9 to access their current account. However, an expert in electronic security systems recommended to the bank management that a more secure system be used. The new system requests the creation of a new password with six alphanumeric characters, now allowing the use of the 26 letters of the alphabet in addition to the digits 0 to 9. In this new system, an uppercase letter is considered different from its lowercase version. The use of other types of characters is prohibited. One way of evaluating a change in the password system is to check for the improvement coefficient, which is the ratio between the number of password possibilities in the new and the old systems. The improvement coefficient of the recommended change is

A $\frac{62^6}{10^6}$
 C $\frac{62! 4!}{10! 56!}$
 E $62^6 - 10^6$

B $\frac{62!}{10!}$
 D $62! - 10!$

Figure 4.13: Problem 3, Bank Password.

The mathematical concept with which most students engaged in this problem was the calculation of the total number of available alphanumeric characters after the change in the system proposed in the problem. That is, in this problem, most students were able to figure that there were 62 different alphanumeric characters to be used in the new system (10 numbers, 26 lower-case letters, and 26 upper-case letters). Among the nine students who engaged conceptually with this problem, four of them engaged *only* with this sum. To a large extent, then, evidence indicates that the problem allowed students access to conceptual reasoning by allowing a simple concept outside of the combinatorial domain (sums) as an entry point.

In contrast to the three problems discussed above, Problem 9, Volleyball Teams, had the lowest level of conceptual engagement across Combinatorics problems (42%), despite being adequately restated by 92% of students. In this problem, I found no evidence of student conceptual engagement through simple combinatorial or non-combinatorial concepts. Student conceptual engagement involved more advanced concepts such as permutations and combinations. The problem is reproduced below along with examples of student quotes.²²

Eight friends, two of them left-handed, decided to make a beach volleyball tournament on their holidays. They need to form four couples for the tournament. None of the couples should have two left-handed players. In how many different ways can these four couples be formed?
 a) 69 b) 70 c) 90 d) 104 e) 105

Figure 4.14: Problem 9, Volleyball Teams

Student 2: It would be a permutation [...] of 8 objects taken 2 at a time. But in my mind, I would pick one pair and the other 6 pairs would be ignored, so I wasn't sure if they would count these values again [...]. I am trying to find a solution that addresses this - one in which I evaluate the other 3 pairs as well, not just picking one pair.

Student 6: I singled out a left-handed person, who can pair up with 6 other people. Then, from that first pair that the left-handed person formed, the next one can only pair up with five more."

Student 9: So, I made 4 combinations. It would be a combination of '8,2' multiplied by '6,2', then '4,2', and '2,2'. Those would be the 4 combinations. This is divided by 4!, which is the number of pairs the tournament requires to be formed. But this is considering that there isn't a strict rule about left-handed people. So, with left-handers, we have to remove one pair."

²² **Student 2:** Seria um arranjo [...] 8 elementos que os 2 a 2. Mas na minha cabeça, a ia pegar uma dupla e as outras 6 duplas vão ser ignoradas então não sabia se eles iam contar esses valores novamente [...]. Eu estou tentando buscar uma solução que passe por isso - que eu, que eu avalie as outras 3 duplas também, além de pegar só uma dupla.

Student 6: Eu separei um canhoto, ele pode fazer dupla com outras 6 pessoas. E aí, a partir dessa primeira dupla que o canhoto fez, o outro só pode fazer com mais cinco.

Student 9: Então seria eu fiz 4 combinações. Seria uma combinação "8,2" multiplicado por "6,2", multiplicado por "4,2" e "2,2". Essas seriam as 4. Isso dividido por 4! que é o número de duplas que o torneio necessita, seja formado. Só que isso é considerando que não há um que não há a regra restritamente do canhoto. Então, com os canhotos, a gente tem que tirar uma dupla

Student 11: Eu fiz primeiro, 7. É, fiz primeiro 7×6 . O 7 era era é tipo as 8 pessoas tirando uma, porque não podia ter 2 canhotos, né? Aí, beleza, eu tirei uma e aí depois eu coloquei as outras 6 pessoas que sobraram.

Student 11: First, I did 7 times 6. The 7 is, like, the 8 people minus one, because you couldn't have 2 left-handers, right? So, okay, I removed one and then I considered the remaining 6 people.

Student conceptual engagement with this problem indicates that, in general, it did not allow for simple concepts to be used as entry points. In fact, in this problem, evidence of conceptual engagement was concentrated on five students who demonstrated high command of Combinatorics concepts throughout the interviews. That is, the students who showed evidence of conceptual engagement in this problem are the exact five students with the highest overall rates of correct response across Combinatorics problems. Accordingly, two students in the provided examples demonstrated familiarity with complex combinatorial concepts in their answers. Student 2 mentions permutations and Student 9 mentions combinations. The other two students in the passages demonstrated a high capacity to exert nonstandard reasoning. Students 6 and 11 tried to create a counting process that fits the problems restriction by leaving one of the left-handed players out. Overall, the problem seems to have a high requirement in terms of prior combinatorial knowledge or problem-solving skills in order to allow for conceptual engagement. This might have prevented students who were familiar only with the multiplication principle (but not with more complex combinatorial concepts), to access the problem.

I conclude from this analysis that Combinatorics problems, in general, do not allow students to engage in conceptual reasoning to the same extent that they allow students to understand what is being asked. While high rates of student understanding seem to happen across varied types of Combinatorial problems, whether conceptual engagement is high or low depends on the specificities of each problem. Across the interviews, problems that allowed for visual

reasoning and for simple mathematical concepts as entry points to solution attempts seem to have allowed more students to engage with conceptual mathematics.

c) Relationship between the absence of verbal cues and difficulty

The two remaining hypotheses drawn from the literature concern the *absence* of something. In this subsection, I analyze the absence of verbal cues, and in the next subsection the absence of powerful methods, and how the two are related to problem difficulty. As it is naturally challenging to exemplify the absence of something, I try to analyze these phenomena through indirect evidence provided in students' interactions with the interview problems.

As pointed out in the literature, distinguishing between appropriate methods for different Combinatorics problems can be challenging to students (e.g., Batanero et al., 1997; Wasserman & Galarza, 2019). An underlying hypothesis for this phenomenon is that the language used in Combinatorics problems frequently does not provide cues into which method is to be used (Daroczy et al., 2015). The difficulties faced by students in figuring out the appropriate combinatorial methods for different problems were indeed evident in some of the interviews and will be analyzed further in this section by focusing on the case of a single problem - Problem 7, Painting Circles. This problem provides an interesting example to look into issues of language and difficulty because it is the most difficult Combinatorics problems in the set and arguably has one of the simplest statements. The problem had the highest rates of incorrect response in the ENEM exam data (89%) and in the interviews (83%). The problem is also the second shortest in the interview set (60 words) and had no pre-identified complex symbol or word.

Before we investigate this case, though, I show that many students explicitly (and spontaneously) expressed that they knew the problem they were working on involved

Combinatorics to some extent. Six students (50%) made such spontaneous references in at least one Combinatorics problem, as illustrated in the following excerpts.²³

Student 5: I'm thinking that it's about Combinatorics. It's a little complicated to... the information, what can I use to do my calculations. (Problem 1, Detective Game.)

Student 9: Well, looking here, I already identified that the theme of this problem is combinatorial analysis. Because it deals with possibilities, in relationship to objects, in relationship to choices. So I already identified that's this. (Problem 1, Detective Game.)

Student 2: Ok, I'm already looking at the options [begins to read the problem out loud]. By the repertoire of problems that I have, I already have an idea of what is probably going to require. Problems about passwords are more or less always the same. (Problem 3, Bank Password.)

While students do not refer to how they figured the problems were in the combinatorial domain, it is possible to identify a common linguistic clue that was present in the majority of combinatorics problems in the list. Six out of eight combinatorial problems asked students to find “how many different ways” something a scenario could be fulfilled, with only slight wording variations. For instance, Problem 4, Renting Movies, asks students “In how many different ways can this customer’s strategy be put into practice?”. Problem 6, Airplane Seats, asks for “the number of different ways to accommodate the family”. One can hypothesize that

²³ Original transcripts:

Student 5, Problem 1, Detective Game: Eu estou achando que é análise combinatória. Tá um pouco complicado de... as informações, o que que eu posso usar pra estar fazendo no meu cálculo.

Student 9, Problem 1, Detective Game: Tá, já to começando vendo as opções. Vai análise combinatória de novo - fatorial, essas coisas. [reads the problem out loud the initial phrase in the problem statement.] Pelo repertório de questões que eu tenho, eu já tenho uma ideia do que provavelmente pode cobrar nisso - questão de senha é mais ou menos sempre a mesma coisa.

Student 2, Problem 3, Bank Password: Bom olhando aqui, eu já identifiquei que o tema dessa questão de análise combinatória. Por tratar de possibilidades em relação a objetos em relação à escolhas. Então, eu já identifiquei que era isso.

the presence of phrases of this form hints students into figuring out that the problem involves combinatorial ideas.

As students also used similar phrases when restating Combinatorics problems in their own words,²⁴ their presence might have contributed to the high rates of correct restatement across Combinatorics problems (70%) when compared to non-Combinatorics problems (54%). While the small number of problems from other domains in the interviews problem does not allow for an in depth comparison, their statements provide initial evidence of the presence of comparatively more varied question phrases. Problem 2, “Container Volume”, a problem in Geometry, asks “What would happen to the water level if we put an object of volume 2400 cm^3 in the tank?”. Another problem in Geometry, Problem 8, Scouts’ Tent, asks for “The projection of the insect’s path on the plane that contains base ABCD”. Both question structures seem unlikely to be replaceable by more general Geometry problem statements akin the general combinatorial phrase “In how many ways can you...”. Because of that, it seems likely that students would need some degree of command of the specific content assessed in each problem (e.g., volume and projections) to rephrase it correctly.

However, figuring which problems involved combinatorial ideas did not prevent students from getting into the many conceptual and procedural pitfalls described in the prior literature. Students frequently demonstrated difficulties in disentangling the particular combinatorial operations involved in Combinatorics problems. I illustrate this phenomenon with an analysis of

²⁴ That is the case, for example of student quotes provided in p.87, in which students explain what is asked in Problem 6, Airplane Seats.

Problem 7, Painting Circles - which was introduced in the previous subsection. In this problem, even though most students could explain what the problem was asking in coherent ways (I marked 92% of students' descriptions of the problem as valid), only one student was able to provide a correct answer accompanied by a coherent mathematical explanation of the method used to solve it. It is thus worthwhile to inspect what sorts of barriers students faced in solving this problem and what prevented them from providing correct answers even though they understood what was being asked.

Problem 7, Painting Circles, has a small numerical answer and thus, in principle, can be solved by listing all cases - but, in practice, no student attempted to do so. Analytical solutions to the problem can be very nuanced, involving combinations (e.g., some cases will involve choosing 2 out of 3 colors), the addition principle (e.g., using *exactly two colors* or using *exactly three colors* are mutually exclusive scenarios), and/or the multiplication principle (e.g., one can start by holding a color constant at a given position). I identified two major categories of mistakes in students' attempts to solve the problem in face of these nuances. I first discuss these two types of mistakes and then revisit the problem to analyze how they may be connected to the language used in its statement.

The first type of mistake was to straightforwardly attempt to use a method that did not fit the problem. Four students did that by trying to use the multiplication principle, as illustrated in the following excerpts.²⁵

²⁵ Original transcripts:

Student 3: *Aí, eu fiz... Eu não fiz uma conta tão grande assim, eu só fiz... É o número de de lápis de cor vezes o número de segmentos que aí deu 12. Espero que esteja certo. [Beleza. E por quê?] E foi a B. [É, e qual foi esse raciocínio de fazer... por que que você fez um número de lápis de cor vezes os segmentos?] Porque aí você ia fazer o número de os lápis e quantas vezes eles poderiam estar nos 4 segmentos, entendeu? Acho que é isso.*

Student 5: *Ele está pedindo de quantas maneiras é na criança pode pintar. Ele só tem 3 cores. E que os demais, tem que ter cores diferentes. Dessa forma, você... bom eu fiz 3 vezes o 4 que são segmentos, que poderia dar 12 maneiras, que é a questão (B).*

Student 3: So, I did... I didn't do that big of a calculation, I just did... It's the number of colored pencils times the number of segments, which gave me 12. I hope it's right. [Interviewer: Okay. And why?] And it was [answer choice] B. [Interviewer: Okay, and what was the reasoning behind doing... why did you do the number of colored pencils times the segments?] Because then you were going to do the number of pencils and how many times could they be in the 4 segments, you understand? I think that's it.

Student 5: He is asking how many ways the child can paint. There are only 3 colors. And for the others, they have to have different colors. This way, you... well, I did 3 times 4, which is the number of segments, which would give 12 ways, which is answer choice B.

Student 7: Yeah, I don't know if it would be right, but I think... Actually, I don't think it's right, but I want to speak within some logic, so I would put, I don't know, 12. [Interviewer: Can you tell me how you got to these 12?] Because of that, that multiplication thing, you know? Anyway, the number of pencils and the quantities.

Student 11: I would go with [answer choice] B, for some calculations that I was doing here, just simple, equal to 3 times 4, anyway. [Interviewer: Okay, and why the 3 times 4?] Yes, the 4 dots, the 4 circles, right? And the 3 different colored pencils.

In the excerpts above, students try to use the multiplication principle in a way that did not properly fit the problem. Their interaction with the problem is somewhat straightforward, in the sense that they test out whether a formula fits one of the possible answer choices and then move on - the four student interactions provided in the excerpts fit my coding criteria for strictly procedural attempts to solve the problem²⁶. This perception is expressed by Student 7, who indicates that they believe their own procedure is incorrect. A rushed interaction with the

Student 7: *É, eu não sei se estaria certo, mas acho que... Na verdade, eu não acho que esteja certo, mas eu quero falar dentro de alguma lógica, então eu colocaria, sei lá, 12. [E conta como que você chegou nessas 12? Por aquela, aquele negócio de multiplicar, sabe? Enfim, a quantidade de lápis e as quantidades.*

Student 11: *Eu iria na (B) por uma contas que eu ia fazer, simples e está 3 vezes 4 mesmo. [Beleza, e porque os 3 vezes 4?] É, as 4 bolinhas, os 4 círculos né, e os 3 lápis de cor diferente.*

²⁶ In the case of Student 7, a second attempt to solve the problem was coded as a conceptual interaction.

problem can also be inferred to some extent from student 3's statement that they "didn't do that big of a calculation" and "hope[s] it's right".

Two factors may have influenced the students to go in this direction. First, the multiplication principle could be straightforwardly applied to previous problems with success. Especially in Problem 1, Detective Game, the approach attempted by these students would likely have worked out fine. Problem 3, Renting Movies, and Problem 6, Airplane Seats, could also be solved using only the multiplication principle - albeit in these cases, it requires an accounting of the details of the problem. By having worked on these three problems prior to the current problem, students might have been wrongly cued into using the multiplication principle. Also, this problem had a strong detractor answer choice. The specific numeric outcome of the straightforward (and incorrect) use of the multiplication principle was one of the answer choices (answer choice B, 12), alluding students to a "confirmation" if they attempted to go this route.

Another type of error observed in Problem 7 was more nuanced. Some students were able to correctly disentangle parts of the problem before committing mistakes. These cases provide a full illustration of what the coexistence of accessibility and difficulty may look like in students' interactions with Combinatorics problems. While students understood what the problem asked and arguably engaged in valuable combinatorial reasoning in their solution attempts, they ended up choosing an incorrect answer to the problem. The following excerpt illustrates on such case.²⁷

²⁷ Original transcripts:

Student 6: Aí, esse essa questão, por ser números pequenos, acho que dá para fazer de um jeito um pouco mais instintivo. Então, eu parti de que na A ela tinha 3 possibilidades de cores. Como ela começou com A, [então] na B ela só teria 2 [possibilidades], porque ela já não pode por mais da mesma cor que A. Na C, ela também tem 2 [possibilidades], porque não tem problema ela repetir a cor da A. Mas ela não pode repetir a cor da B, então sobram 2 opções de cores. E na D é a mesma coisa, ela não pode repetir a cor da A, nem da C. Então.. Ela tem, é... Ela tem uma cor, porque, se ela já tem 2 ligadas ali. Seguindo, pensando que ela fez a ABCD, um então na D ela só tem uma cor que sobra porque ela já usou uma na A e outra na C, então sobra uma. Então, no total dá 12. Que seria B.

Student 6: *So, this question, because the numbers are small, I think it can be done in a slightly more instinctive way. So, I assumed that in [circle] A there were 3 color possibilities. As she started with A, in B she would only have 2 [possibilities] because she can no longer have the same color as A. In C, she also has 2 [possibilities] because there is no problem for her to repeat the color of A. But she cannot repeat the color of B, so there are 2 color options left. And in D it's the same thing, the color that's in A can't repeat, nor [the color that is in] C. So... there is... There is one color [left] because, there are already 2 [painted circles] connected there. Assuming that she did ABCD, then in D she only has one color left because she already used one in A and another in C, so there is only one left. So, in total it gives 12. Which would be [answer choice B].*

Student 6 was able to delve into consistent mathematical sensemaking by analyzing what happens if one paints the circles sequentially, from A to D. The student makes a mistake only in analyzing how many options would be available to paint circle D. This is, in fact, a critical mathematical nuance of this problem. Using the sequential painting procedure proposed by the student, the number of available colors to paint D depends on whether the color used in C was the same as in A. If A and C are the same color, then any of the two remaining colors can be used in D. But if A and C are different colors, then there is only one color left to paint D (the color that was neither used in A nor C).

Except for this challenging nuance, the way the student used the multiplication principle in the problem was internally consistent. The student consistently justified why each number was multiplied in the sequence they proposed (three ways to paint A, times two ways to paint B, and so forth). Arguably, the student engaged in valuable mathematical thinking while performing this solution attempt. The proposed solution, albeit incorrect, demonstrated consistent combinatorial sense-making through a justified counting process. A nuance in the problem (the two separate cases that needed to be accounted for in the case of Circle D) held the student back from arriving at a correct answer.

Student 8 started with a similar reasoning - sequentially computing the possibilities of colors available for each circle - and then also incurred in a similar mistake as Student 6.²⁸

Student 8: At first, I will multiply 3, because in A I can use any of the 3 colors that I wish. I can choose any of the 3, so I have 3 options. In B, I have 2 [options] because [if I choose yellow in A] then I can use blue or red there, but I can't use yellow. So that's times 2. Then in D it's also 2. So, I will multiply in this way. Three times two...

The main difference in Student 8's reasoning in comparison to Student 6's reasoning is that instead of considering only the case in which D has one available option (A and C are painted different colors), Student 8 considered only the case in which D has two colors available (A and C are painted the same color). Student 8 also arrived at an incorrect response (24). Similarly to the previously discussed excerpt, there is evidence of valuable combinatorial reasoning in this procedure, which nonetheless led the student to the incorrect answer choice due to not considering a nuanced part of the problem.

As mentioned, Problem 7, Painting Circles, comprised one of the simplest problem statements in the list. It is the second shortest problem and had no pre-identified complex vocabulary. It is worthwhile to investigate whether this relatively simple problem statement offers any cues about the critical mathematical nuance that held Students 6 and Student 8 from arriving at a correct answer. While part of the discussed nuance relates to the provided rule that circles "connected by a segment must have different colors", it becomes evident only in connection with the details in the image. To correctly implement a sequential counting

²⁸ Original transcripts:

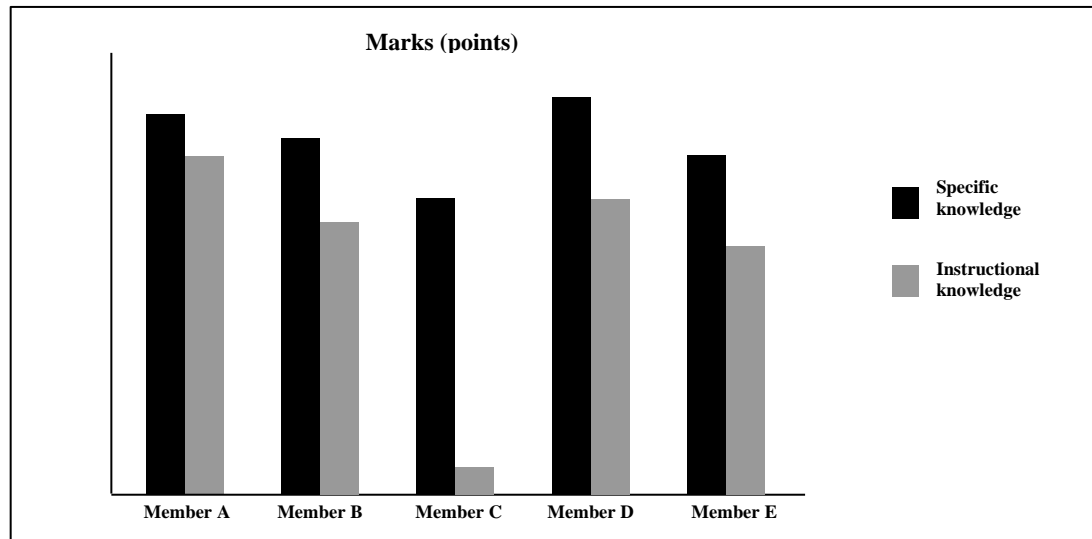
Student 8: Na primeira, vou multiplicar 3, porque no A eu posso colocar as 3 cores que eu quiser. Eu posso escolher uma das 3, então eu tenho 3 opções. Na B eu já tenho 2 porque [se na A eu tiver colocado amarelo] eu posso colocar azul e vermelho, mas não posso colocar amarelo. Então é vezes 2. Aí na D também são duas. Enfim, eu vou multiplicar nesse sentido aí. Três vezes dois...

procedure, the student must make sense of the rule restricting connected circles to be of different colors, and effectively account for its interaction with the shape of the provided drawing. In this sense, the written statement of the problem does not provide a direct hint to students about the most challenging aspect of the problem.

To provide an additional perspective on the issue of linguistic clues and student difficulties in solving problems, a comparison can be made with Problem 5, Hiring Process, a problem in Data Analysis. Differently from the Painting Circles problem, in the Hiring Process problem, most students who understood what the problem asked could correctly answer *and* provide coherent explanations for their solutions. Eight students provided explanations of what the problem asked that I coded as valid, and five of those responded to the problem correctly and with valid justification, as per my coding. In comparison, only one student coherently justified a correct choice in the Painting Circles problem, among 11 who showed evidence of having understood what was being asked.

Prior to the interviews, I identified that Problem 5, Hiring Process, had one potentially challenging mathematical term, *average*. Seven students provided coherent explanations of what the term “average” means, and among them were the five students who answered the problem correctly and provided compelling justifications. So, while this mathematics-specific term could, in principle, prevent some students from understanding what the problem asked, I hypothesize that it may also have provided other students with an indication of the mathematics needed to solve the problem. That is, for some students, the presence of a mathematics-specific term “average” may have been helpful as a hint to the core procedure needed to solve the problem, which mainly involved calculating and then taking the difference between two averages.

A candidate for a faculty position participated in a selection process in which the recruiting panel was composed of five members. The marks received by the candidate in the process are shown in the graph.



Each member of the panel attributed two marks to the candidate, one relative to *specific knowledge* of the area of expertise and the other to *instructional knowledge*, and the final summary score of the candidate was given by the arithmetic mean of all the marks attributed by the members of the panel. Using a new criterion, this evaluation panel decided to discard the highest and lowest grades attributed to the candidate. The new summary score, relative to the old summary score, is

- a) 0.25 points higher.
- b) 1.00 point higher.
- c) 1.00 point smaller.
- d) 1.25 points higher.
- e) 2.00 points lower.

Figure 4.15: Problem 5, Hiring Process.

The example of Problem 5, Hiring Process, shows that linguistic cues to appropriate solutions methods can be provided through mathematics-specific vocabulary, implying that the *absence of linguistic cues* about adequate solution methods and the *simplicity of vocabulary* may actually be interconnected. These two characteristics can be interpreted as flip sides of the same coin. Simpler problem statements (simplicity of vocabulary) may tend to afford less cues about appropriate solution paths (absence of linguistic cues).

d) Relationship between the absence of powerful methods and difficulty: Strategic guessing

One aspect related to the absence of powerful methods overlaps with the previous discussion about the absence of verbal cues. We delved into examples of students trying to use a combinatorial method in a way that did not fit the problem they were working on. In Problem 3, Painting Circles, some students tried to use the multiplication principle in a straightforward manner that did not help them solve the problem. One can interpret this as evidence that students are trying to use *some* method (in the example discussed, the multiplication method) *as if* it were a powerful method. In this sense, the examples provided in the previous section can be interpreted as evidence that aligns with the claim in literature that points out to the absence of powerful methods in Combinatorics as an explanatory factor for difficulty (e.g., Kapur, 1970; Lockwood et al., 2020).

In this subsection, I focus on a different phenomenon related to solution methods and problem difficulty. I call this phenomenon, which I observed in the data and had not anticipated prior to the interviews, *answer-based guessing*. Answer-based guessing comprises situations in which students explicitly analyze the characteristics of the answer choices to try to make an informed guess about which one is correct. For instance, in Problem 3, Bank Password, students would look for answer choices that included the numbers provided in the problem statement and guess that such answers were more likely to be correct. Table 4.10 presents the frequency of student interactions with problems that included answer-based guessing attempts, as per my coding, and reproduces an example of an answer choice available in each problem.


	<i>Domain</i>	<i>Answer Choice example</i>	<i>Answer-based guessing</i>
<i>Pair of problems A</i>			
<i>Problem 1- Detective game</i>	<i>Combinatorics</i>	<i>a) 10 more students than possible different answers.</i>	<i>8%</i>
<i>Problem 2- Container volume</i>	<i>Geometry</i>	<i>a) The level would rise 0.2 cm, making the water 20.2 cm high.</i>	<i>33%</i>
<i>Pair of problems B</i>			
<i>Problem 3 - Bank password</i>	<i>Combinatorics</i>	<i>a) $\frac{62^6}{10^6}$</i>	<i>33%</i>
<i>Problem 5- Teacher hiring</i>	<i>Data analysis</i>	<i>a) 0.25 points higher.</i>	<i>0%</i>
<i>Pair of problems C</i>			
<i>Problem 7- Painting circles</i>	<i>Combinatorics</i>	<i>a) 6</i>	<i>8%</i>
<i>Problem 8- Scouts' tent</i>	<i>Geometry</i>	<i>a) </i>	<i>0%</i>
<i>Pair of problems D</i>			
<i>Problem 12 - Paths on map</i>	<i>Combinatorics</i>	<i>a) 4</i>	<i>8%</i>
<i>Problem 11- Motorbike trip</i>	<i>Algebra</i>	<i>a) 13.</i>	<i>0%</i>
<i>Other Combinatorics problems</i>			
<i>Problem 4 - Renting movies</i>	<i>Combinatorics</i>	<i>a) $20 \times 8! + (3!)^2$</i>	<i>83%</i>
<i>Problem 6 - Aitplane seats</i>	<i>Combinatorics</i>	<i>a) $\frac{9!}{2!}$</i>	<i>50%</i>
<i>Problem 9 - Volleyball teams</i>	<i>Combinatorics</i>	<i>a) 69</i>	<i>33%</i>
<i>Problem 10- Toy train</i>	<i>Combinatorics</i>	<i>a) $C_{12}^4 \times C_{12}^3 \times C_{12}^3 \times C_{12}^2$</i>	<i>50%</i>

Table 4.10: Answer-based guessing

As shown in Table 4.10, answer-based guessing was pronounced in problems that contained mathematical expressions in their answer choices (Problems, 4,6, and 10), which are all problems in Combinatorics. It seems likely thus, that the more information is provided in the answer choices, the more students try to use that information in strategic guessing. While in

some cases students combined a dual strategy by advancing on some depth of mathematical sensemaking and only then resorted to answer-based guessing, this approach was also used as the single justification provided by students for their answer choices. An example of answer-based guessing in which a student is very explicit about the use of such strategy in Problem 4, Renting Movies, is provided below, along with the problem statement.²⁹

A customer of a video rental store has the habit of renting two movies at a time. When he returns them, he always picks up two other movies - and so on. He learned that the video store received some new releases, 8 action movies, 5 comedy movies, and 3 drama movies, so he established a strategy to see all 16 releases. Initially, at each visit, he will rent an action movie and a comedy movie. When comedy possibilities are exhausted, the customer will rent an action movie and a drama movie, until all releases are watched. No movie will be rented more than once.

In how many different ways can this customer's strategy be put into practice?

a)	$20 \times 8! + (3!)^2$	c)	$\frac{8! \times 5! \times 3!}{2^8}$	e)	$\frac{16!}{2^8}$
b)	$8! \times 5! \times 3!$	d)	$\frac{8! \times 5! \times 3!}{2^2}$		

Figure 4.16. Problem 4, Renting Movies.

Student 7: Look, honestly, this thing, this topic is not something I know a lot about. So, again, I would guess using certain logic, let's say. I really don't know much about where to begin, so I would pick [answer choice] C or D. [...] I don't know, it's because it just looks more right. Like, I wouldn't understand why you would do 20 times 8, with this symbol, plus 3 [in answer choice A]. I wouldn't understand that. So, I know that probably the calculations in this area are closer to [answer choices] C or D [...] I would guess C, because "power 8" also repeats itself in E. (Problem 4, incorrect answer.)

²⁹ Original quote:

Student 7: *Olha, sinceramente, esse negócio.. essa área não é algo que eu sei muito. Então novamente, eu chutaria com uma certa lógica, diremos assim, é... um eu realmente não sei muito para mim começar, então eu colocaria ou (C) ou (D). [...] Não sei, é porque simplesmente me parece mais certo, tipo, eu não entenderia porque você faria, por exemplo, 20×8 a com esse símbolo, mais 3 [item (A)] eu não não entenderia. Assim, eu sei que provavelmente as contas de dessa matéria se estão mais próximas de um (C) ou de um (D). [...] Eu chutaria a (C). porque o elevado a 8 está na (E) também.*

In the passage, the student explains that, due to an absence of prior knowledge in the content area assessed in Problem 4, they decided to focus on the answers to try to find out which answer choice was the most likely. The student initially selects two answer choices and then uses a criterion - based solely on the information contained in the answer choices themselves - to pick one of the two. As explained in the passage, answer choice C was chosen by the student because it contained a term that was also present in another answer choice (E).

On the same problem, Student 4 arrived at the correct answer choice using a similar strategy:³⁰

Student 4: I was like, between [answer choices] E and B. Like, in my mind the most correct, from what I got from the text, would be B. But [answer choice] C, has this 2 and the power 8, that I... I think it means the possibilities [Interviewer: But what did you start by saying you thought that B looked right?] Because, like, it makes sense, 8 times 5 times 3 is what gives you the overall possibilities. (Problem 4, correct answer.)

In the passage, Student 4 seems to combine an initial attempt to make sense of the mathematics of the problem with a subsequent exploration of information from the answer choices. This combination of mathematical sensemaking and strategic choice was present in other cases as well. In the passage below, Student 7, whose answer-based guessing in Problem 4 led to a wrong answer, applies the same reasoning to Problem 10, Toy Train.³¹

³⁰ Original transcripts:

Student 4, Problem 4 (correct answer): Eu fiquei tipo, em dúvida entre a E e a B. Tipo na minha cabeça assim o mais certo assim, pelo que eu entendi do texto, seria a B. Só que a C, tipo, está lá, 2, que é o casalzinho e tipo e elevado a 8 que eu acho, eu acho que significa as possibilidades [Interviewer: Mas por que você começou falando que você achava que a B provavelmente a princípio, e parecia certo?] Porque, tipo, faz sentido, tipo 8 vezes 5×3 é que tipo daria as possibilidades no geral.

³¹ Original transcripts:

Student 7, Problem 10 (correct answer): Primeiro que ia ser a entre A, B e E. Aí, agora eu estou tentando nichar. Acho que agora está entre a A ou a E. Acho que é a E. [Interviewer: Tá, por quê?] Então, eu... Eu parti da ideia daquele negócio que eu falei de quando você utiliza das cadeiras, lembra que eu falei que quando um senta o outro não pode ocupar? Então, eu fui por eu só tava entre a questão do 12 embaixo, né? Que aqui está todos na A né? Todos são 12 embaixo e aqui [na E] tem essa variação que você vai tirando anterior. Então se seguiu mais ou menos

Student 7: First, I was between A, B and E. So, now I'm trying to niche it down. I think now it's between A or E. I think it's E. [Interviewer: Okay, why?] So, I... I started from the idea of that thing I talked about before [in the airplane problem]. When you use chairs, remember, I said that when one sits, the other can't occupy it? So, I went this way because I was in doubt in regard to the [number] 12 at the bottom [of the answer choices], right? It's in all [terms of answer choice] A, right? They are all 12 at the bottom. And here [in answer choice E], there is this variation that you take from the previous one. So this is the idea I more or less followed, I took it, like, from the previous one. If you look at it, like, 8 means 12 minus 4, which is 8. Hence 8 minus 3 is 5. So on and so forth. 5 minus 3 gives 2. And so on. (Problem 10, correct answer.)

This time, the student chooses the correct answer to the problem. Interestingly, this student was not able to explain the meaning of the term “combinations”, as discussed in a previous subsection. Using an answer-based guessing strategy, though, the student found a pattern in the answer choices and interpreted it correctly (i.e. in the notation used in the answer choices, $C(n, k)$, term k represented the number of wagons “still available” to be painted with each color).

The emergent phenomenon of answer-based guessing is particularly interesting because it can potentially shed light on three stylized facts previously highlighted in this dissertation about the connection between accessibility and difficulty. First, it was noted that, overall, students in the interviews were on average more correct in their answer choices in Combinatorics problems (when compared to non-Combinatorics problems), but this higher level of correct responses was not accompanied by a similar level of coherence in students' explanations about how they solved the problems. That is, many students could pick the right answer but could not explain the mathematics behind their choices. Answer-based guessing provides a possible explanation for at

essa ideia, eu fui tirando, tipo, do anterior. Que se você for ver, tipo o 8 quer dizer os 12 menos 4, que dá 8. Daí 8 menos 3 dá 5. Assim por aí vai. O 5 menos 3 dá 2. E aí vai.
[...][...]

least part of this phenomenon, as students applied this strategy more frequently in Combinatorics than in non-Combinatorics problems.

Second, answer-based guessing may also shed light on a possible mechanism behind the results obtained in my analysis of Research Question 1. In that quantitative analysis, evidence pointed that Combinatorics problems in ENEM present, in general, lower school and racial gaps than non-Combinatorics problems. It may be the case that part of the reason why gaps are narrower in Combinatorics problems is due to a higher presence of formulas in its answer choices, which leads to more answer-based guessing attempts. Third, answer-based guessing also provides a possible explanation for the fact that in the quantitative analysis, the presence of mathematical symbols was correlated with lower achievement gaps. If mathematical notation in ENEM problems is primarily concentrated in the answer choices (rather than on problem statements), it might be the case that answer-based guessing is more prevalent in the problems that were coded as containing complex notation. The coding scheme used in the quantitative analysis did not discern, though, whether symbols were present in problems' statements or answer choices.

Answering Research Question 3

To conclude this chapter, I briefly provide interpretations of how evidence from the interviews helps shed light on Research Question 3.

What are some of the ways in which students access problems in Combinatorics? The literature review in Chapter 2 highlights the idea that Combinatorics problems may be particularly accessible because they do not require extensive mathematics-specific terminology to be stated. The qualitative analysis of student interviews provided examples showing that

mathematics-specific terms may also be an important factor in determining accessibility levels *across Combinatorics problems*. Also, the qualitative evidence presented in this chapter supports the idea that students used the context of Combinatorics problems to understand what problems asked, in more useful ways than they did in non-Combinatorics problems. Finally, it was also shown that student conceptual engagement levels can vary significantly across Combinatorics problems. In particular, problems that allow for entry points that combine visual representations and simple mathematics concepts (e.g., the multiplication principle) seem to be associated with higher conceptual engagement.

What are some of the ways in which students fail to provide correct answers to problems in Combinatorics? In line with the results from previous studies, this chapter has presented evidence that one of the difficulties students face in solving Combinatorics problems is choosing an appropriate combinatorial method. The analysis of students' interactions with Combinatorics problems unveiled examples of students attempting to use a Combinatorial method – the multiplication principle – “as if” it were a powerful method. These results align with the idea that Combinatorics lacks “powerful methods”, and always require “ingenuity” (Kapur, 1970). The examples discussed in this chapter also shed light on how Combinatorics problems may include nuances not directly referenced in the problem's statement, making it challenging for students to answer problems correctly.

What is the relationship between difficulty and accessibility in students' interactions with Combinatorics problems? The discussions in this chapter raised the hypothesis that the absence of mathematics-specific vocabulary and symbols may be in fact be a “two-edged sword”. While it may make it easier for some students to understand what problems ask, it may also prevent them from being clued into specific combinatorial methods that are adequate for

solving the problem. Another relationship between accessibility and difficulty was revealed in student's interactions with the available answer choices in closed-ended problems. When closed-ended Combinatorics problems have formulas in their answer choices, they seem to allow students to use answer-based guessing, a strategic pathway to try to figure out the correct answer choice without necessarily understanding its underlying mathematical content. Students seemed to be more inclined to use this strategy in problems with higher difficulty levels. This phenomenon can provide a partial explanatory hypothesis for the patterns observed in the ENEM data, in which Combinatorics problems have been shown to have higher difficulty levels and smaller gaps than problems in other domains. It might be the case that Combinatorics problems in the ENEM exam have answer choices that are more prone to be used in answer-based guessing, due to a higher presence of formulas. If the efficacy of the answer-based guessing strategy is equally distributed between public and private school students, for example, this phenomenon may explain part of the reduced private-public school gap observed in Combinatorics problems.

Chapter V: Conclusion

This chapter concludes my dissertation. Initially, I recapitulate how the findings of the study address its proposed research questions. Next, I provide a comprehensive discussion highlighting the contributions of this study. Finally, I acknowledge the study's limitations and suggest directions for further research.

Summary of Results

This dissertation aimed to delve into the paradoxical character of Combinatorics, presenting it as a domain in Mathematics that is both difficult and accessible. To explore these dimensions, a mixed-methods approach was employed. Quantitative data was gathered from the ENEM, a comprehensive nationwide assessment in Brazil, to compare difficulty and accessibility levels between Combinatorics and non-Combinatorics problems. Additionally, qualitative methods, specifically task-based interviews with students from disadvantaged school backgrounds in Brazil, were conducted. These interviews sought to offer deeper insights into the strategies students employ, their understanding, and their level of engagement with Combinatorics and non-Combinatorics problems.

Research Question 1. Using ENEM data on exams taken by thousands of high-school students across 12 years, this study showed that *difficulty* and *accessibility* levels, as measured by the proxy indicators used in this study, were higher in Combinatorics than in non-Combinatorics problems. These results were significant and stable under different empirical specifications. Interview findings aligned in part with these results. In alignment with the quantitative findings, student interviews revealed that while the combinatorial domain is notably accessible in the sense that it allows students with varied backgrounds to understand what problems ask, this accessibility does not necessarily translate into students consistently arriving at correct solutions.

In contrast with the quantitative findings, Combinatorics problems did not present higher levels of student conceptual engagement with mathematics in student interviews, when compared to non-Combinatorics problems.

Research Question 2. Also using a quantitative procedure, this study explored possible predictors of difficulty and accessibility levels across Combinatorics and non-Combinatorics problems. Overall, results indicated that the small set of quantifiable characteristics of problems used (length of the problem statement and presence of mathematical notation), did not explain away the differences found in difficulty and accessibility between Combinatorics and non-Combinatorics problems. Nevertheless, these linguistic characteristics have shown significant coefficients of association with difficulty and accessibility levels. In particular, the presence of mathematics notation was associated with higher difficulty levels and with lower achievement gaps.

Interview findings further explored the relationship between language, difficulty, and accessibility. Students' interactions with Combinatorics problems indicated that the absence of mathematics-specific language and symbols may be a two-edged sword. On the one hand, the use of simpler language may render problems more accessible by allowing students with varied mathematical backgrounds to understand what is being asked. On the other hand, the absence of mathematics-specific language can make problems more difficult by omitting written cues about what combinatorial method might be used to solve those problems.

In the interviews, many students analyzed the available answers choices of problems to make strategic guesses about which of them were most likely correct. I called this phenomenon answer-based guessing. Students seemed to use the answer-based guessing strategy more frequently in problems with more information in the answer choices. More specifically, problems

with formulas in their answer choices have shown to be particularly prone to answer-based guessing. This pattern aligns with the quantitative results indicating a positive association between presence of mathematics notation and accessibility (i.e. lower achievement gaps).

Research Question 3. The qualitative portion of this study aimed to unveil some of the ways in which students interact with Combinatorics and non-Combinatorics problems in the ENEM exam. First, it was shown that students may leverage the context of Combinatorics problems to make sense of their underlying mathematics. Using the context of Combinatorics problems, students were able to summarize them in compact ways. Students also relied on context to communicate the meaning of important mathematical ideas of Combinatorics problems. While these results align, in general, with the literature in showing that Combinatorics problems can be particularly accessible, the specific importance of context for student understanding of what problems ask highlights a novel aspect of this phenomenon. It was also shown that the level of conceptual engagement among students can differ greatly depending on the type of Combinatorics problems they encounter. Specifically, problems that incorporate visual elements and basic mathematical concepts, such as the multiplication principle, tended to be linked with increased conceptual engagement levels. Finally, the results from the interviews confirmed previous findings that students struggle to select the right combinatorial approach in Combinatorics problems. Students often attempted to use the multiplication principle incorrectly, suggesting a lack of “powerful methods” in Combinatorics, as proposed by Kapur (1970).

Discussion and Contributions

Arguably, the most innovative results from this dissertation are comparisons of *difficulty* and *accessibility* levels in Combinatorics and non-Combinatorics problems. This study

contributes to the literature by adding large-scale empirical evidence to a growing consensus about the dual nature – comparably difficult and accessible – of the combinatorial domain in secondary education. While accessibility and difficulty in Combinatorics problems had been previously documented in qualitative and small-scale quantitative studies, the literature still lacked the type of comparative, large-scale analysis presented in this dissertation.

Taken from a broader perspective, these results indicate the possibility that *difficulty* and *accessibility* might be domain-specific traits in mathematics problems. That is, the evidence that difficulty and accessibility were higher in Combinatorics problems than in non-Combinatorics problems in the ENEM exam can be interpreted as suggestive that other domains may also show differing patterns of difficulty and accessibility. The fact that students' interactions with Combinatorics problems aligned with some of the most prevalent explanations provided in the literature for the dual nature of Combinatorics provides additional evidence that difficulty and accessibility may be related to the ontology of mathematical domains. More specifically, this study has provided evidence that Combinatorics problem statements with less complex vocabulary tend to be more accessible (more students understand what the problems ask). It was also shown that students could not resort to “powerful methods” to solve Combinatorics problems and that this was associated with higher difficulty levels. Both these characteristics have been referred to in the literature as specific traits of the Combinatorial domain. Other mathematical domains might possess different characteristics, which can influence accessibility and difficulty in varied ways.

Some of the results obtained in this study can expand the current theoretical discussions about possible mechanisms influencing accessibility and difficulty levels in Combinatorics problems, while not directly contradicting them. To summarize these adaptations, Figure 5.1

represents the updated hypotheses about the determinants of accessibility and difficulty in Combinatorics problems.

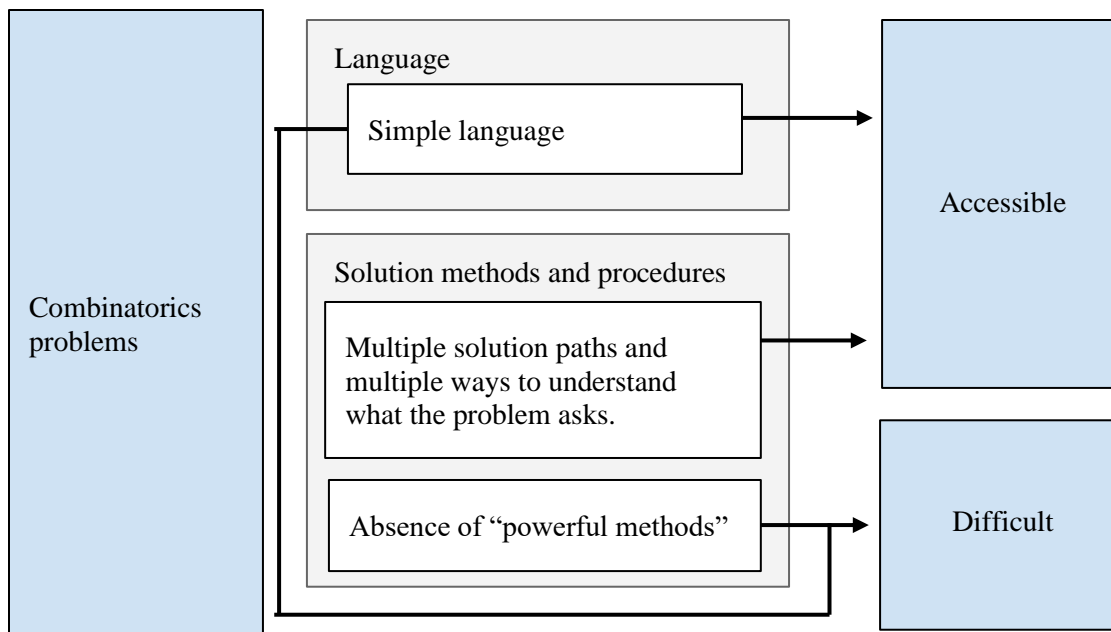


Figure 5.1: Schematic representation of updated hypotheses

The updated framework presents three main changes in comparison to the hypothesis initially drawn from the literature. First, the category “absence of verbal cues” is incorporated into the framework as part of the category “simple language”. As illustrated in the discussions of Problem 7, Painting Circles, and Problem 5, Hiring Process, I hypothesize that the absence of mathematics-specific concepts and symbols may prevent Combinatorics problems from providing cues about which combinatorial methods are most appropriate to be used. That is, the specific way in which many Combinatorics problems omit verbal cues about appropriate solution methods is by not using complex mathematics language.

Second, student interviews provided evidence that “simple language” and “absence of powerful methods” have a combined influence on making Combinatorics problems particularly difficult. When the simplicity of the language used in problem statements prevents students from

being hinted into appropriate solution methods, students may try to use a Combinatorial method “as if” it were a powerful method that could solve a wide variety of problems. The cases discussed in the previous chapter have shown that students tried to use the multiplication method in ways that did not fit the problems they were working on.

Third, I expand the category “multiple solution paths” to include “multiple ways to understand what problems ask”. The analysis of interviews showed that students effectively used the context to understand what Combinatorics problems asks. This use of real-world examples in Combinatorics suggests that students draw upon their diverse experiences to understand the questions posed in this domain.

Beyond the initial hypotheses framework drawn from the literature, this study has also uncovered some unanticipated phenomena related to accessibility and difficulty levels of Combinatorics problems. Evidence was presented that there can be wide variability in access levels of problems within the Combinatorial domain. Specifically, results from the interviews indicate that although students seem to understand a wide range of Combinatorial problems, the level of conceptual engagement varies based on the problem's specifics. In the set of problems used for the interviews, those that permitted visual reasoning and employed straightforward mathematical concepts as starting points for solutions appeared to encourage more engagement with conceptual mathematics.

An additional unpredicted contribution from this study emerged from the use of a mixed methods approach. Many authors have highlighted that this comprehensive research method can be beneficial in providing multiple perspectives on educational phenomena (Creswell & Poth, 2017; Ponce & Pagán-Maldonado, 2015). The combined results from the quantitative and qualitative analyses in this dissertation provide a clear example of this perspective. Using the

same types of problems present in the quantitative analysis, the qualitative section of this study has uncovered the unpredicted phenomenon of *answer-based guessing*, which modulates how the quantitative results should be interpreted. In the interviews, students were more frequently inclined to use this strategy in problems that had mathematical expressions in their answer choices. If Combinatorics problems tend to include mathematical expressions in their answer choices, students might be using answer-based guessing more frequently when working on problems in this domain than when working on problems in other domains. The hypothesis that answer-based guessing may be one of the contributing factors to the narrower gaps observed in Combinatorics problems also aligns with an empirical result that the presence of mathematical symbols in ENEM problems was correlated with lower achievement gaps. If symbols in ENEM questions are concentrated in the answer choices (rather than on problem statements), students might be more likely to engage in answer-based guessing in problems that include such symbols.

How this phenomenon affects the estimates of accessibility in Combinatorics and non-Combinatorics depends how often and how well students from different socioeconomic backgrounds use this strategy in the ENEM exam. If, answer-based guessing is a strategy that it can be successfully used by students with varied backgrounds - then the difference in accessibility levels found in the quantitative analysis (e.g., smaller school gaps in Combinatorics problems) may in part result from this phenomenon. That is, if answer-based guessing is evenly used by public and private school students, it can work towards “leveling the playfield” between these two groups of students. If this is the case, then the results found in this study likely overstate the “true” difference in accessibility levels between Combinatorics and non-Combinatorics problems. If, on the other hand, private school students are more prone to using answer-based guessing than public school students (suppose, for example, that private schools

are more focused on “teaching to the test” and might explicitly teach students how to draw information from answer choices), then the results found in this study underestimate the “true” difference between accessibility levels in Combinatorics and non-Combinatorics problems.

The evidence that Combinatorics problems vary in how they afford conceptual engagement has important implications for the teaching and learning of Combinatorics. It suggests that, for Combinatorics lessons to be fully accessible to students with varied backgrounds, problems must be carefully selected. Data from student interviews indicate that problems that allow for visual forms of reasoning and include simple mathematics concepts as entry points may be particularly suitable for conceptual engagement. Teachers interested in introducing Combinatorics problems to their students with the goal of providing equitable access to conceptual mathematics can prioritize problems with these characteristics.

Another result that can inform teaching is that many students who failed to provide correct answers to Combinatorics problems in the interviews actually engaged in valuable mathematical thinking while working on such problems. This phenomenon was discussed in detail in the case of Problem 7, Painting Circles. A first implication of this result for equitable teaching of Combinatorics is that classrooms should be more focused on the process rather than on the final answers provided by students. Teachers can, for example, provide feedback based on the full extent of students’ written works, rather than solely on their final answers. A second implication for teaching relates to the result that some students were held back from correct answers by nuances of problems that were not directly hinted in the problems’ statements. This phenomenon was also investigated in detail in Problem 7, Painting Circles. Teachers may benefit from being aware of which problems contain such challenging nuances to decide on how to use them in their courses. Teachers interested in developing students’ problem-solving skills may

decide to provide these types of problems to students without providing any additional hints. Teachers interested in developing preliminary Combinatorial ideas, on the other hand, might hold back from using such problems or may explicitly hint students into their most challenging nuances. Finally, teachers interested in both may allow students to initially work on the problem on their own - and then explicitly discuss the nuance with the whole class before allowing additional time for students to review their solutions.

Limitations and Pathways for Research

While the quantitative section of this study is arguably reliable due to the large sample size it encompasses, it is also limited by being based on a single exam. The ENEM exam has a specific format (multiple choice questions), a typical style of problems (highly contextualized), and a specific scope (the Brazilian secondary mathematics curriculum). It is essential to have this context in mind to interpret results accordingly. Further research with different formats of Combinatorial problems could inform whether the results obtained in the context of ENEM are consistent across different Combinatorics problem formats and styles (e.g., open-ended questions, decontextualized questions).

Also, the results obtained in comparing Combinatorics and non-Combinatorics problems should be interpreted as associated with the whole scope of non-combinatorial domains assessed in ENEM. In this sense, this study sheds little light on how accessibility and difficulty levels of Combinatorics problems compare specifically to Algebra, Geometry, or any other single domain of mathematics. Further research could address this issue by comparing difficulty and accessibility levels between problems in Combinatorics and problems in other specific domains.

This study's quantitative results on accessibility are limited by using proxy indicators for accessibility (achievement gaps). While achievement gaps have been shown to be particularly narrow in Combinatorics problems of the ENEM exam, the interviews suggest that, to some extent, this phenomenon may be related to answer-based guessing. This casts doubt on the interpretation that the proxy indicators used in the study adequately capture accessibility levels of problems in the ENEM exam. Further investigations could improve our understanding of the extent to which narrower gaps in Combinatorics are associated with high levels of accessibility, and to what extent they are associated with answer-based guessing. An initial step towards this would be to investigate how more detailed indicators of linguistic complexity than the ones used in this study correlate with achievement gaps. The qualitative analysis in this study suggests, for example, that the presence of formulas in problems may influence students' interactions differently if the formulas are present in the problems' statements or in their answer choices - but this distinction was not analyzed quantitatively in this study.

Also, further qualitative investigation focused specifically on the phenomenon of answer-based guessing could bring important lessons about the role played by answer choices in student mathematical reasoning. While this study has documented some examples of answer-based guessing in which students did not delve much into the mathematics of problems, it was also shown that students used answer choices as leverage to mathematical understanding. Future research focused specifically on answer-based guessing could shed light on the relationship between problem statements, problem answer choices, and student engagement with mathematics.

References

- Allison, P. D. (1999). *Multiple Regression: A Primer*. Pine Forge.
- Anderson, I., van Asch, B., & van Lint, J. (2004). Discrete mathematics in the high school curriculum. *Zentralblatt Für Didaktik Der Mathematik*, 36(3), 105–116. <https://doi.org/10.1007/bf02652778>
- Andreescu, T., & Feng, Z. (2003). A path to combinatorics for undergraduates: Counting strategies. Birkhauser Boston. Andreescu, T., & Feng, Z. (2003). *A path to combinatorics for undergraduates: Counting strategies* (2004th ed.). Birkhauser Boston.
- Annin, S. A., & Lai, K. S. (2010). Common errors in counting problems. *The Mathematics Teacher*, 103(6), 402–409. <https://doi.org/10.5951/mt.103.6.0402>
- Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.
- Armor, D. J., Marks, G. N., & Malatinszky, A. (2018). The impact of school SES on student achievement: Evidence from U.S. statewide achievement data. *Educational Evaluation and Policy Analysis*, 40(4), 613–630. <https://doi.org/10.3102/0162373718787917>
- Avalos, M. A., Medina, E., & Secada, W. G. (2018). Reading Mathematics Problems. In *Language, Literacy, and Learning in the STEM Disciplines* (1st Edition, pp. 55–78). Routledge.
- Becker, K. L., & Arends-Kuenning, M. P. (2020). Analysis of Math Test Score Gap Between White and Non-White Students in Brazilian Public Schools: SAEB 2015. *The Review of Black Political Economy*, 47(2), 138–158. <https://doi.org/10.1177/0034644619879381>
- Batanero, C., Navarro-Pelayo, V., & Godino, J. (1997). Effect of the implicit combinatorial model on combinatorial reasoning in secondary school pupils. *Educational Studies in Mathematics*, 32(1), 181-199. <https://doi.org/10.1023/A:1002954428327>
- Blanton, M., Stephens, A., Knuth, E., Gardiner, A. M., Isler, I & Kim, S.. (2015). The Development of Children's Algebraic Thinking: The Impact of a Comprehensive Early Algebra Intervention in Third Grade. *Journal for Research in Mathematics Education*, 46(1), 39–87. <https://doi.org/10.5951/jresmetheduc.46.1.0039>
- Blanton, M. L., & Kaput, J. J. (2005). Characterizing a Classroom Practice That Promotes Algebraic Reasoning. *Journal for Research in Mathematics Education*, 36(5), 412–446. <http://www.jstor.org/stable/30034944>
- Boaler, J., & Greeno, J. (2000). Identity, agency, and knowing in mathematics worlds. In J. Boaler (Ed.), *Multiple perspectives on mathematics teaching and learning*. 171–200. Westport, CT: Ablex.

- Bülbul, B. Ö., & Güler, M. (2021). Can Geometry Achievement and Geometric Habits of Mind Be Improved Online? Reflections From a Computer-Aided Intervention. *Journal of Educational Technology Systems*, 49(3), 376–398.
<https://doi.org/10.1177/0047239520965234>
- Burger, W. F., & Shaughnessy, J. M. (1986). Characterizing the van hiele levels of development in geometry. *Journal for Research in Mathematics Education*, 17(1), 31.
<https://doi.org/10.2307/749317>
- Cai, J., Morris, A., Hohensee, C., Hwang, S., Robison, V., Cirillo, M., Kramer, S. L., & Hiebert, J. (2019). Posing significant research questions. *Journal for Research in Mathematics Education*, 50(2), 114–120. <https://doi.org/10.5951/jresmetheduc.50.2.0114>
- Cameron, P. J. (1994). *Combinatorics: Topics, Techniques, Algorithms*. Cambridge University Press.
- Carraher, D. W. & Schliemann, A. D. (2007). Early algebra and algebraic reasoning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning*, 669-705. NCTM.
- Carraher, D. W. , Schliemann, A. D. (2020). Early Algebra Teaching and Learning. In: Lerman, S. (eds) *Encyclopedia of Mathematics Education*. Springer. https://doi.org/10.1007/978-3-030-15789-0_54
- Cavalcanti, T., Guimaraes, J., & Sampaio, B. (2010). Barriers to skill acquisition in Brazil: Public and private school students' performance in a public university entrance exam. *The Quarterly Review of Economics and Finance: Journal of the Midwest Economics Association*, 50(4), 395–407. <https://doi.org/10.1016/j.qref.2010.08.001>
- Chancel, L., Piketty, T., Saez, E., & Zucman, G. (2022). World Inequality Report 2022, World Inequality Lab. www.wir2022.wid.world
- Charles, R. I.,(2005). Big ideas and understandings as the foundation for elementary and middle school mathematics. *Journal of Mathematics Education*, 7(3), 9-24.
- Charmaz, K. (2014). *Constructing grounded theory* (2nd ed.). SAGE Publications.
- Cohen, E. G., & Lotan, R. A. (Eds.). (1997). *Working for equity in heterogeneous classrooms: Sociological theory in practice*. Teachers' College Press.
- Creswell, J. W., & Poth, C. N. (2017). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications.
- Cuoco, A., Paul Goldenberg, E., & Mark, J. (1996). Habits of mind: An organizing principle for mathematics curricula. *The Journal of Mathematical Behavior*, 15(4), 375–402.
[https://doi.org/10.1016/s0732-3123\(96\)90023-1](https://doi.org/10.1016/s0732-3123(96)90023-1)

- Darling-Hammond, L. (2014). Closing the achievement gap: A systemic view. In J. V. Clark (Ed.), *Closing the Achievement Gap from an International Perspective* (pp. 7–20). Springer Netherlands.
- Daroczy, G., Wolska, M., Meurers, W. D., & Nuerk, H.-C. (2015). Word problems: a review of linguistic and numerical factors contributing to their difficulty. *Frontiers in Psychology*, *06*, 1–13. <https://doi.org/10.3389/fpsyg.2015.00348>
- Debeer, D., & Janssen, R. (2013). Modeling Item-Position Effects Within an IRT Framework. *Journal of Educational Measurement*, *50*(2), 164–185. <http://www.jstor.org/stable/24018105>
- DeBellis, V. A., & Rosenstein, J. G. (2004). Discrete mathematics in primary and secondary schools in the United States. *Zentralblatt Für Didaktik Der Mathematik*, *36*(2), 46–55. <https://doi.org/10.1007/bf02655758>
- Driscoll, M. (1999). *Fostering algebraic thinking: A guide for teachers, grades 6-10*. Heinemann Educational Books.
- Driscoll, M., Dimatteo, R., Nikula, J., & Egan, M. (2017). *Fostering geometric thinking: A guide for teachers, grades 5-10*. Heinemann Educational Books.
- Dubois, J. G. (1984). Une systématique des configurations combinatoires simples. *Educational Studies in Mathematics*, *15*(1), 37–57.
- English, L.D. (1991) Young children's combinatoric strategies. *Educational Studies in Mathematics*, *22*(5), 451–474. <https://doi.org/10.1007/BF00367908>
- English, L. D. (1993). Children's strategies for solving two- and three-dimensional combinatorial problems. *Journal for Research in Mathematics Education*, *24*(3), 255–273.
- English, L. D. (2005). Combinatorics and the development of children's combinatorial reasoning. In G.A. Jones (Eds) *Exploring probability in school* (pp. 121-141). Springer. https://doi.org/10.1007/0-387-24530-8_6
- English, L. D., & Watson, J. M. (2016). Development of Probabilistic Understanding in Fourth Grade. *Journal for Research in Mathematics Education*, *47*(1), 28–62. <https://doi.org/10.5951/jresmetheduc.47.1.0028>
- Erickson, M. (2010). Change for a dollar, change for a million. *Math Horizons*, *17*(3), 22–25. <https://doi.org/10.4169/194762110x489224>
- Goldin, G. A. (2004). Problem solving heuristics, affect, and discrete mathematics. *Zentralblatt Für Didaktik Der Mathematik*, *36*(2), 56–60. <https://doi.org/10.1007/bf02655759>

- Gowers, T., Barrow-Green, J., & Imre (Eds.). (2008). *The Princeton companion to mathematics*. Princeton University Press.
- Gutiérrez, R. (2007). Context matters: Equity, success, and the future of mathematics education. In T. Lamberg & L. R. Wiest (Eds.), *Proceedings of the 29th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 1-18). Stateline (Lake Tahoe), Nevada. University of Nevada, Reno.
- Gutiérrez, A., Jaime, A., & Fortuny, J. M. (1991). An Alternative Paradigm to Evaluate the Acquisition of the van Hiele Levels. *Journal for Research in Mathematics Education*, 22(3), 237–251. <https://doi.org/10.2307/749076>
- Gutiérrez, R. (2008). Research commentary: A gap-gazing fetish in mathematics education? Problematizing research on the achievement gap. *Journal for Research in Mathematics Education*, 39(4), 357–364. <https://doi.org/10.5951/jresmetheduc.39.4.0357>
- Hadar, N., & Hadass, R. (1981). The road to solving a combinatorial problem is strewn with pitfalls. *Educational Studies in Mathematics*, 12(4), 435–443. <https://doi.org/10.1007/bf00308141>
- Halani (2013). *Students' Ways of Thinking about Combinatorics Solution Set*. Doctoral Dissertation. Arizona State University. <https://keep.lib.asu.edu/items/151965/view>
- Hardy, M. (1993). Regression with dummy variables. *Quantitative Applications in Social Sciences* (93). Sage.
- Hegedus, S. J., & Moreno-Armella, L. (Eds.). (2009). Transforming mathematics education through the use of dynamic mathematics technologies [Special Issue]. *ZDM – The International Journal on Mathematics Education*, 41(4).
- Hollebrands, K. F. (2007). The role of a dynamic software program for geometry in the strategies high school mathematics students employ. *Journal for research in mathematics education*, 38(2), 164-192. <https://www.jstor.org/stable/30034955>
- Hurdle, Z., Warshauer, M., & White, A. (2016). The place and purpose of combinatorics. *The Mathematics Teacher*, 110(3), 216–221. <https://doi.org/10.5951/mathteacher.110.3.0216>
- IBGE (2021). *Rendimento todas as fontes 2020*. Instituto Brasileiro de Geografia e Estatística. Retrieved from https://biblioteca.ibge.gov.br/visualizacao/livros/liv101880_informativo.pdf.
- Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (2009). Matriz de Referência do ENEM. INEP. https://download.inep.gov.br/download/enem/matriz_referencia.pdf

- Kapur, J. N. (1970). Combinatorial analysis and school mathematics. *Educational Studies in Mathematics*, 3(1), 111–127.
- Kingston, N. M., & Dorans, N. J. (1984). Item Location Effects and Their Implications for IRT Equating and Adaptive Testing. *Applied Psychological Measurement*, 8(2), 147–154. <https://doi.org/10.1177/014662168400800202>
- Lamanna, L., Gea, M., & Batanero, C. (2022). Secondary school students' strategies in solving combination problems. *Twelfth Congress of the European Society for Research in Mathematics Education (CERME12)*, 1-9. Hal-03751834. <https://hal.archives-ouvertes.fr/hal-03751834>
- Lockwood, E. (2011). Student connections among counting problems: an exploration using actor-oriented transfer. *Educational Studies in Mathematics*, 78(3), 307–322. <https://doi.org/10.1007/s10649-011-9320-7>
- Lockwood, E., Wasserman, N. H., & McGuffey, W. (2018). Classifying combinations: Investigating undergraduate students' responses to different categories of combination problems. *International Journal of Research in Undergraduate Mathematics Education*, 4(2), 305–322. <https://doi.org/10.1007/s40753-018-0073-x>
- Lockwood, E., Wasserman, N. H., & Tillema, E. S. (2020). A case for combinatorics: A research commentary. *The Journal of Mathematical Behavior*, 59(100783), 100783. <https://doi.org/10.1016/j.jmathb.2020.100783>
- Lockwood, E., & Gibson, B. R. (2016). Combinatorial tasks and outcome listing: Examining productive listing among undergraduate students. *Educational Studies in Mathematics*, 91(2), 247–270. <https://doi.org/10.1007/s10649-015-9664-5>
- Lockwood, E., & Purdy, B. (2019). Two undergraduate students' reinvention of the multiplication principle. *Journal for Research in Mathematics Education*, 50(3), 225–267. <https://doi.org/10.5951/jresematheduc.50.3.0225>
- Lubienski, S. T. (2008). On “gap gazing” in mathematics education: The need for gaps analyses. *Journal for Research in Mathematics Education*, 39(4), 350–356. <http://www.jstor.org/stable/40539301>
- Madeira, R. A., & Rangel, M. A. (2014). Racial achievement gaps in another America: Discussing schooling outcomes and affirmative action in Brazil. In *Closing the Achievement Gap from an International Perspective* (pp. 127–160). Springer Netherlands. https://doi.org/10.1007/978-94-007-4357-1_7
- Maher, C. A., Powell, A. B., & Uptegrove, E. B. (2010). *Combinatorics and reasoning: Representing, Justifying and Building Isomorphisms*. Springer.

- Maher, C.A., Sigley, R. (2020). Task-Based Interviews in Mathematics Education. In: Lerman, S. (eds) *Encyclopedia of Mathematics Education*. Springer, Cham.
https://doi.org/10.1007/978-3-030-15789-0_147
- Mamona-Downs, J., & Martin Downs. (2004). Realization of Techniques in Problem Solving: The Construction of Bijections for Enumeration Tasks. *Educational Studies in Mathematics*, 56(2/3), 235–253. <http://www.jstor.org/stable/4150283>
- Martiniello, M. (2009). Linguistic complexity, schematic representations, and differential item functioning for English language learners in math tests. *Educational Assessment*, 14(3–4), 160–179. <https://doi.org/10.1080/10627190903422906>
- Martino, A. M. (1992). *Elementary students' construction of mathematical knowledge: Analysis by profile*. Unpublished doctoral dissertation, Rutgers University.
- Mashiach Eizenberg, M., & Zaslavsky, O. (2004). Students' verification strategies for combinatorial problems. *Mathematical Thinking and Learning*, 6(1), 15–36.
https://doi.org/10.1207/s15327833mtl0601_2
- Mathison, S. (1988). Why Triangulate? *Educational Researcher*, 17(2), 13.
<https://doi.org/10.2307/1174583>
- McGraw, R., Lubienski, S. T., & Strutchens, M. E. (2006). A Closer Look at Gender in NAEP Mathematics Achievement and Affect Data: Intersections with Achievement, Race/Ethnicity, and Socioeconomic Status. *Journal for Research in Mathematics Education*, 37(2), 129–150. <http://www.jstor.org/stable/30034845>
- Mello, U. (2022). Centralized admissions, affirmative action, and access of low-income students to higher education. *American Economic Journal: Economic Policy*, 14(3), 166-197.
<https://doi.org/10.1257/pol.20190639>
- National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. Authors.
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for Mathematics*. Authors.
<https://learning.ccsso.org/wp-content/uploads/2022/11/ADA-Compliant-Math-Standards.pdf>
- Navarro-Pelayo, V. (1994). Estructura de los problemas combinatorios simples y del razonamiento combinatorio en alumnos de secundaria. Tesis Doctoral. Universidad de Granada.
- OECD (2020). *PISA 2018 Results (Volume V): Effective Policies, Successful Schools*. OECD Publishing. <https://doi.org/10.1787/ca768d40-en>

- Oliveira, P. R. de, Belluzzo, W., & Pazello, E. T. (2013). The public–private test score gap in Brazil. *Economics of Education Review*, 35, 120–133. <https://doi.org/10.1016/j.econedurev.2013.04.003>
- Polya, G. (1945). *How to solve it: A new aspect of mathematical method*. Princeton University Press.
- Ponce, O. A., & Pagán-Maldonado, N. (2015). Mixed methods research in education: Capturing the complexity of the profession. *International journal of educational excellence*, 1(1), 111-135.
- Piaget, J., & Inhelder, B. (1951/1975). *The origin of the idea of chance in children*. W.W. Norton. (Original work published 1951).
- Sandefur, J., Lockwood, E., Hart, E., & Greefrath, G. (2022). Teaching and learning discrete mathematics. *ZDM: The International Journal on Mathematics Education*, 54(4), 753–775. <https://doi.org/10.1007/s11858-022-01399-7>
- Schoenfeld, A. H. (1985). *Mathematical problem solving*. Elsevier.
- Schoenfeld, A. H. (2020). Mathematical practices, in theory and practice. *ZDM: The International Journal on Mathematics Education*, 52(6), 1163–1175. <https://doi.org/10.1007/s11858-020-01162-w>
- Silva, M. (2016) A combinatória: abordagem em documentos oficiais, em resultados de pesquisas e em livros didáticos do Ensino Fundamental (Doctoral dissertation). Recife, UFPE. <https://repositorio.ufpe.br/handle/123456789/17752>
- Sinclair, N., & Robutti, O. (2012). Technology and the role of proof: The case of dynamic geometry. In *Third International Handbook of Mathematics Education* (pp. 571–596). Springer New York.
- Souza, W., Oliveira, V., & Anegues, A. C. (2018). Background familiar e desempenho escolar: uma abordagem não paramétrica. *Pesquisa e Planejamento Economico*, 48(2), 133–162. IPEA.
- Stein, M. K., Kaufman, J. H., Sherman, M., & Hillen, A. F. (2011). Algebra: A challenge at the Crossroads of policy and practice. *Review of Educational Research*, 81(4), 453–492. <https://doi.org/10.3102/0034654311423025>
- Tillema, E. S. (2018). An investigation of 6th graders’ solutions of Cartesian product problems and representation of these problems using arrays. *The Journal of Mathematical Behavior*, 52, 1–20. <https://doi.org/10.1016/j.jmathb.2018.03.009>

- Tillema, E. S. (2020). Students' solution of arrangement problems and their connection to Cartesian product problems. *Mathematical Thinking and Learning*, 22(1), 23–55.
<https://doi.org/10.1080/10986065.2019.1608618>
- Tillema, E., & Gatzka, A. (2016). A Quantitative And Combinatorial Approach To Non-linear Meanings Of Multiplication. *For the Learning of Mathematics*, 36(2), 26-33.
<http://www.jstor.org/stable/44382709>
- Tucker, A. (2012). *Applied Combinatorics* (6th ed.). John Wiley & Sons.
- van Hiele, P. M. (1986) *Structure and insight: a theory of mathematics education*. Academic.
- Wasserman, N. H., & Galarza, P. (2019). Conceptualizing and justifying sets of outcomes with combination problems. *Investigations in Mathematics Learning*, 11(2), 83–102.
<https://doi.org/10.1080/19477503.2017.1392208>
- Weigand, H.-G. (2004). Sequences—Basic elements for discrete mathematics. *Zentralblatt Für Didaktik Der Mathematik*, 36(3), 91–97. <https://doi.org/10.1007/bf02652776>
- Winther, J. M., & Golgher, A. B. (2010). Uma investigação sobre a aplicação de bônus adicional como política de ação afirmativa na Universidade Federal de Minas Gerais (UFMG). *Revista Brasileira de Estudos de Populacao*, 27(2), 333–359.
<https://doi.org/10.1590/s0102-30982010000200007>
- Žakelj, A., & Klančar, A. (2022). The role of visual representations in geometry learning. *European Journal of Educational Research*, 11(3), 1393-1411.
<https://doi.org/10.12973/eu-jer.11.3.1393>

Appendix A: Interview Protocol

Protocol Title: Combinatorics Problems:
Evidence From Large-scale Assessments and Student Interviews

Principal Researcher: Fernando Carnauba, Teachers College
+55 11 930 714 321, fa2529@tc.columbia.edu

(1) Before getting started

Read the following introduction to the interviewee:

Thank you for being here. This Zoom call interview will be video-recorded. I will notify you when the video recorder is started and stopped. I am interested in how people think when they work on different problems. So, we will try to go over 12 different mathematics problems. The interview can last at most one hour and a half, and we can take a break in the middle if you wish. So, at many points, I may ask that we should move on to the next problem.

Have in mind that we are partners in this - just by trying out those problems, you will help me understand how people think through a problem. So it is very ok not to be able to finish a problem. I personally find all the problems in the list challenging, so don't be surprised if you also find them to require some deep thinking. I will ask questions like: in your own words, what the problem is asking? And I will also ask you to "think out loud" while you solve it. I will also ask you if there is any word or symbol that you don't know in the problem. And sometimes I'll just be silent to let you take your time to think. I cannot help you with the problems, except by doing one of three things: suggesting that you re-read the problem; or that you explain what you are thinking; or that you try to use drawings or schemes to help you think.

At any point, if you don't want to continue anymore, you can choose to just stop. There is no problem if you are not feeling well and don't want the interview to remain going. If you stay for at least 60 minutes, you will be fully compensated with a bookstore voucher of R\$120. If you decide to leave before 60 minutes, then you will not receive any compensation. Please get a pen and paper. Writing really helps us think and organize our thoughts, so feel free to really use it. Your annotations also help me understand what you are thinking, so don't erase them or throw them away. Do you have any questions before we begin?

(2) For each problem

Time.

The maximum allotted time for each problem is 7 minutes. If the interviewee finishes the problem before that or does not want to work on it anymore, go on to the next problem.

Getting started.

Ask the following questions after the student has read the problem on her own:

Do you think you understood what the problem is asking?

Please tell me in your own words: what is this problem asking?

Without actually answering the problem, tell me what does a solution to this problem look like?

Can you please tell me what the word [specific vocabulary identified for the problem] means?

Can you please tell me what the symbol [specific symbol identified for the problem] means?

Are there any other words or symbols that you are not sure about?

Working on the problem.

Tell the student to start working on the problem and remind her to verbalize how she is thinking about her solution. Try to stay silent as long as you can.

If the student gets stuck.

Try providing the student with one of the following suggestions, according to what you think will better help the person get ‘unstuck’:

Maybe you can read the problem again (e.g., if you notice she got some information wrong).

Please explain to me what you were trying (e.g., if you think she can build off what she started).

Try using paper and pencil (e.g., if you think she can benefit from that - for example, if it is hard to keep track of every case).

Appendix B: Interview Problems

Problem 1 - Detective Game (Combinatorics)

A school principal invited 280 third-year students to participate in a game. Suppose there are 5 objects and 6 characters in a house with 9 rooms; one of the characters hides one of the objects in one of the rooms of the house. The game's objective is to guess which object was hidden by which character and in which room of the house. All students decided to participate. Each time, a student is drawn randomly to try to answer the challenge. The answers must always be different from all of the previous ones, and the same student cannot be drawn more than once. If the student's answer is correct, the student is declared the winner, and the game ends. The principal knows that a student will eventually provide the correct answer because there are

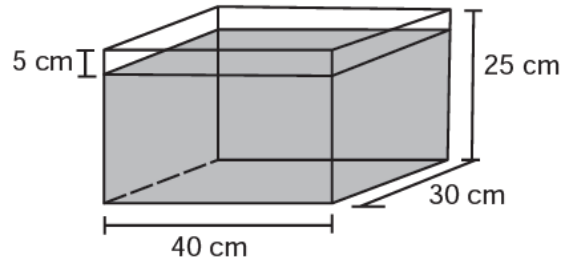
- A) 10 more students than possible different answers.
- B) 20 more students than possible different answers.
- C) 119 more students than possible different answers.
- D) 260 more students than possible different answers.
- E) 270 more students than possible different answers.

Comment: This problem provides a context for the straightforward use of the multiplication principle.

Solution example: Each of the 5 objects can be combined with each of the 6 characters, forming $5 \times 6 = 30$ ordered pairs (*object, character*). Each object-character pair can be combined with each room, forming $30 \times 9 = 270$ ordered triplets (*object, character, room*). With 280 students playing the game, we conclude that there are 10 more students than possible different answers (Answer A).

Problem 2 - Container Volume (Geometry)

During their manufacture, some objects need to go through a cooling process. To do this, a factory uses a cooling tank, as shown in the figure.



What would happen to the water level if we put an object of volume 2400 cm^3 in the tank?

- a) The level would rise 0.2 cm, making the water 20.2 cm high.
- b) The level would rise 1 cm, making the water 21 cm high.
- c) The level would rise 2 cm, making the water 22 cm high.
- d) The level would rise 8 cm, causing the water to overflow.
- e) The level would rise 20 cm, causing the water to overflow.

Comment: Correctly answering this problem seems to depend on the conceptual understanding of volume - just knowing the formula for the volume of rectangular prisms may not be enough.

Solution example: The base of the cooling tank has an area of $40 \times 30 = 1,200 \text{ cm}^2$. Since the tank is a prism, an increase of 1 cm in the height of the water is thus associated with $1,200 \text{ cm}^3$ of “filled-up volume” in the tank. The object that will be put inside is twice as much as this value ($2,400 \text{ cm}^3 = 2 \times 1,200 \text{ cm}^3$), so the water level will rise by 2 cm (Answer C).

Problem 3 - Bank Password (Combinatorics)

A bank requires its customers to create a six-digit password using numbers 0 to 9 to access their current account. However, an expert in electronic security systems recommended to the bank management that a more secure system be used. The new system requests the creation of a new password with six alphanumeric characters, now allowing the use of the 26 letters of the alphabet in addition to the digits 0 to 9. In this new system, an uppercase letter is considered different from its lowercase version. The use of other types of characters is prohibited. One way of evaluating a change in the password system is to check for the improvement coefficient, which is the ratio between the number of password possibilities in the new and the old systems. The improvement coefficient of the recommended change is

- A** $\frac{62^6}{10^6}$ **C** $\frac{62! 4!}{10! 56!}$ **E** $62^6 - 10^6$
B $\frac{62!}{10!}$ **D** $62! - 10!$

Comment: As was the case of problem 1, this problem can also be solved with the multiplication principle. In this case, though, there are two factors that can make the problem more challenging. First, there is the use of combinatorial notation in the answers, with some of them including the symbol for factorials (“!”). Second, the student must know the meaning of *ratio* to solve the problem.

Solution example: Before the change, there were 10 possible digits (0, 1, 2, 3, 4, 5, 6, 7, 8, and 9) to be chosen in each of the six positions in the password (first digit, second digit, third digit... sixth digit). There were thus $10 \times 10 \times 10 \times 10 \times 10 \times 10 = 10^6$ possible keywords. After the change, there were 62 possible characters (any of the 26 lowercase letters or any of the 26 uppercase letters or any of the 10 digits) to be chosen in each of the six positions in the password. There were thus $62 \times 62 \times 62 \times 62 \times 62 \times 62 = 62^6$ possible keywords. The ratio between the new number of available passwords and the old number of available passwords is thus $62^6/10^6$ (Answer A).

Problem 4 - Renting Movies (Combinatorics)

A customer of a video rental store has the habit of renting two movies at a time. When he returns them, he always picks up two other movies - and so on. He learned that the video store received some new releases, 8 action movies, 5 comedy movies, and 3 drama movies, so he established a strategy to see all 16 releases. Initially, at each visit, he will rent an action movie and a comedy movie. When comedy possibilities are exhausted, the customer will rent an action movie and a drama movie, until all releases are watched. No movie will be rented more than once.

In how many different ways can this customer's strategy be put into practice?

- a) $20 \times 8! + (3!)^2$ c) $\frac{8! \times 5! \times 3!}{2^8}$ e) $\frac{16!}{2^8}$
b) $8! \times 5! \times 3!$ d) $\frac{8! \times 5! \times 3!}{2^2}$

Comment: As was the case of problems 1 and 3, this problem can also be solved with the multiplication principle. In this case, though, the problem nuances in how the combinatorial operations should be considered. A first nuance regards the restriction of no repetition of movies. The second nuance is the restriction imposed on the order of the pairs of movies rented - first an action and a comedy movie, and then an action and a drama movie. The factorial notation in this problem might also be a challenge for some students.

Solution example: For the first visit, the customer can choose any of the 8 action movies and any of the 5 comedy movies. There are thus 8×5 to do so. On the second visit, there are 7 action movies left to choose from and 4 comedy movies. There are thus 7×5 ways to choose among the remaining movies in the second visit. One can go on with the same reasoning for the first five visits, leading to the following "partial result" to the problem:

Number of ways to pick movies on the first 5 visits to the store: $(8 \times 5)(7 \times 4)(6 \times 3)(5 \times 2)(4 \times 1)$

On the sixth visit to the store, there are no comedy movies left, so the customer starts to rent an action and a drama movie. There will be 3 action movies left, along with three 3 movies. Also following the same reasoning as before, we arrive at a second "partial result":

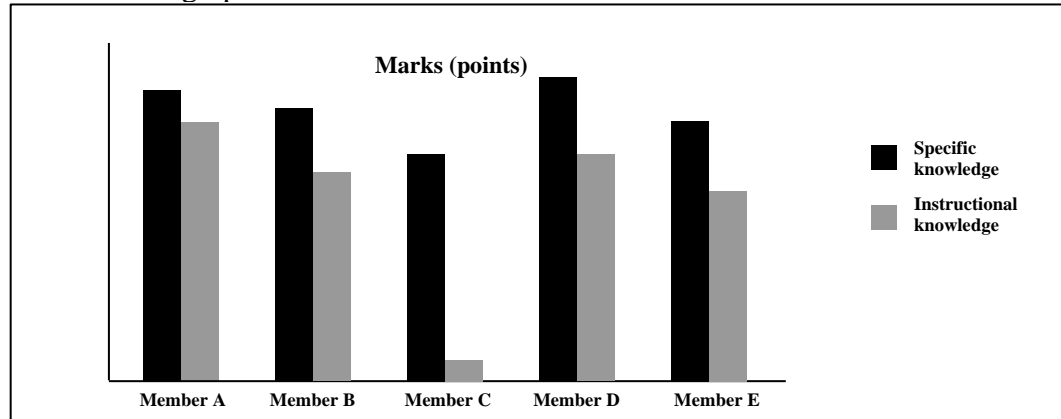
Number of ways to pick movies on the last 3 visits to the store: $(3 \times 3)(2 \times 2)(1 \times 1)$

Considering all 8 visits to the store gives us the following final result for the number of ways in which the customer can implement the proposed strategy:

$$(8 \times 5)(7 \times 4)(6 \times 3)(5 \times 2)(4 \times 1)(3 \times 3)(2 \times 2)(1 \times 1) = 8! \times 5! \times 3! \text{ (Answer B).}$$

Problem 5 - Hiring Process (Data Analysis)

A candidate for a faculty position participated in a selection process in which the recruiting panel was composed of five members. The marks received by the candidate in the process are shown in the graph.



Each member of the panel attributed two marks to the candidate, one relative to *specific knowledge* of the area of expertise and the other to *instructional knowledge*, and the final summary score of the candidate was given by the arithmetic mean of all the marks attributed by the members of the panel. Using a new criterion, this evaluation panel decided to discard the highest and lowest grades attributed to the candidate. The new summary score, relative to the old summary score, is

- (a) 0.25 points higher. (c) 1.00 point smaller. (e) 2.00 points lower.
(b) 1.00 point higher. (d) 1.25 points higher.

Comment: This question uses both graph interpretation and the concept of average value. While manually computing the values can be cumbersome, it is possible to combine values to simplify the computation (see example below).

Solution example: To facilitate the calculation, one can combine terms in that sum to round tens:

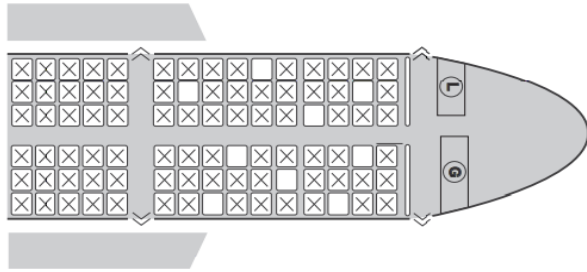
$$\text{“Old” score: } [(18+12)+(16+14)+(17+13)+(1+19)+(14+16)]/10 = [30+30+30+20+30]/10 = 14$$

$$\text{“New” score: } [140-1-19]/8 = 120/8 = 15$$

$$\text{“New” score} - \text{“old” score} = 15 - 14 = 1 \text{ (Answer B).}$$

Problem 6 - Airplane Seats (Combinatorics)

A family of seven adults, after deciding the itinerary of their trip, consulted the website of an airline and found that the flight for the chosen date was almost full. In the figure available on the website, the seats occupied are marked with an X, and the only seats available are those shown in white.



The number of different ways to accommodate the family on that flight is given by

- a) $\frac{9!}{2!}$ c) $7!$ e) $\frac{5!}{4!} \times \frac{4!}{3!}$
 b) $\frac{9!}{7! \times 2!}$ d) $\frac{5!}{2!} \times 4!$

Comment: Although it is possible to solve this problem with reasoning based on the multiplication principle, this problem had a potentially confounding nuance - the two seats that are to remain vacant might be considered distinguishable or indistinguishable, depending on how the counting process is encoded.

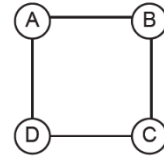
Solution example: Let's assign family members to the seats in increasing order of age. There are 9 distinct seats available for the youngest person. There are 8 seats available for the second-youngest person. We keep doing this until the oldest person (7th to be assigned) has three possible seats to choose from.

Number of ways to assign the seats: $9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 = 9!/2!$ (Answer A).

An alternative route would be to run through the seats and assign either a passenger or one of two empty-seat labels to each of them (e.g. Empty₁, Empty₂). There are 9 possibilities for the first seat (7 passengers and two empty-seat labels), 8 possibilities for the second seat, and so forth. There are thus 9! ways to do this assignment. One would then need to realize that outcomes that differ only in having the two empty-seat labels swapped with each other are equivalent (family members are in the same seats). To take that double counting into account, the previous answer must be divided by 2.

Problem 7 - Painting Circles (Combinatorics)

To stimulate his daughter's reasoning, a father made the following drawing and handed it to the child along with three pencils of different colors. He wants the child to paint only the circles and asks that those connected by a segment must have different colors.

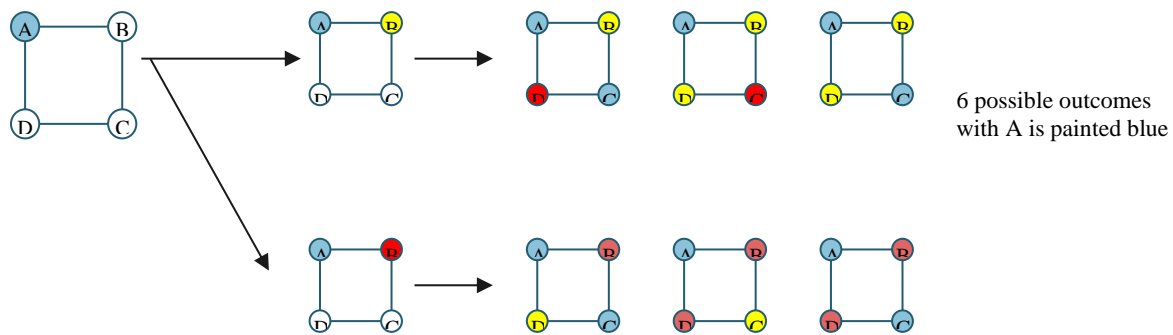


In how many different ways can the child do what her father asks?

- a) 6 b) 12 c) 18 d) 24 e) 7

Comment: This question has a small numerical answer and thus, in principle, can be solved by listing all cases. More analytical solutions to the problem can be very nuanced, involving combinations (e.g., some cases will involve choosing 2 out of 3 colors), the addition principle (e.g., using *exactly two colors* or using *exactly three colors* are mutually exclusive scenarios), and/or the multiplication principle (e.g., one can start by holding a color constant at a given position).

Solution example: Suppose we have the colors blue, yellow, and red. Let's start by figuring out how many possibilities we have if we paint circle A with the blue color.



Now we can realize that we could also have started by painting circle A with the yellow pen, which would give us 6 new coloring schemes. Also, painting circle A red would give us 6 more. We arrived at $6 \times 3 = 18$ possible coloring schemes in total (Answer C).

Problem 8 - Scouts' Tent (Geometry)

A group of Boy Scouts was doing an activity in the park in the city where they live and set up a tent as shown in the photo of Figure 1. Figure 2 shows a schematic representation of the frame of this tent, which has the shape of a straight prism made of metal bars.



Figura 1

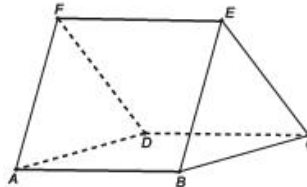


Figura 2

After the frame was assembled, one of the scouts observed an insect moving over them from vertex A towards vertex B, then from there towards vertex E, when finally it made the trajectory from vertex E to C. Assume that all these displacements were made the shortest path between the points. The projection of the insect's path on the plane that contains base ABCD is given by



Comment: This problem seems to require prior knowledge of the geometric concepts of *right prism*, *vertical projection*, and *plane* (that contains a set of given points). The answer options include a distractor that is based on the displacement of the insect based on the “photographic” representation of the tent (answer B).

Solution example: One must notice that the problem states that the tent is a *right prism*. Because of that, the projection of the edges of the tent on plane ABCE has the shape of two rectangles.

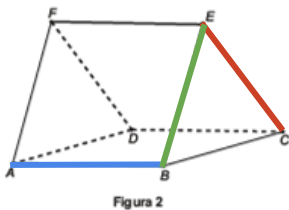
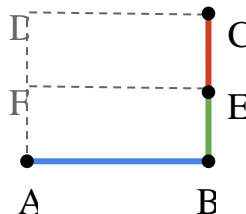


Figura 2



The path taken by the insect thus only uses straight angles and takes the shape of an “inverted L” (Answer E).

Problem 9 - Volleyball Teams (Combinatorics)

Eight friends, two of them left-handed, decided to make a beach volleyball tournament on their holidays. They need to form four couples for the tournament. None of the couples should have two left-handed players. In how many different ways can these four couples be formed?
 a) 69 b) 70 c) 90 d) 104 e) 105

Comment: This seems to be a very challenging problem, with no straightforward solution procedure applicable. While using combinations can be helpful (e.g., “choosing k players among n” to be in a team), there are many complicating factors in this problem. First, there are four teams to be created, so students will need to keep track of a counting process that accounts for all of them. Second, there is the imposition of the rule teams cannot have two left-handed players.

Possible answer: Let's name the two left-hand players L_n and the six right-hand players R_n :

L₁
L₂
R₁
R₂
R₃
R₄
R₅
R₆

We can start by assigning teammates to the left-hand players. For example, L_1 can be assigned with R_1 while L_2 gets assigned with R_2 :

L₁R₁
L₂R₂

There are 6 ways to assign a right-hand player to L_1 and then there are five remaining ways to assign a right-hand player to L_2 . There are thus $6 \times 5 = 30$ ways to assign the two left-hand players to their teams. For each of these 30 ways to assign the left-hand players into teams, we are still left to decide on how to assign the remaining 4 right-hand players into two more teams. In the previous example (L_1R_1, L_2R_2) there would be three possible ways to do this, as shown below.

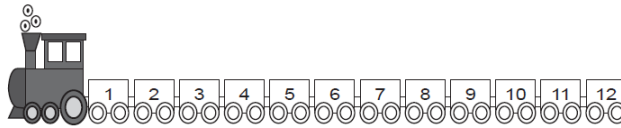
R₃R₄ | R₃R₅ | R₃R₆
R₅R₆ | R₄R₆ | R₄R₅

One way to see that there are only three options is to realize that the only decision that needs to be made is *who will be paired with R_3* , since the remaining two players will forcefully be together in a team). By symmetry, we conclude that there are $30 \times 3 = 90$ ways to create couples for the championship (Answer C).

Problem 10 - Toy Train (Combinatorics)

A company makes toy trains comprised of a black locomotive and 12 wagons numbered 1 to 12. Of the 12 wagons, 4 are painted red, 3 are painted blue, 3 are painted green, and 2 are painted yellow.

Trains are assembled starting with the locomotive and attaching 12 numbered wagons in increasing order, as shown in the image below.



The number of different trains that can be made, as expressed by combinations, is given by

- A $C_{12}^4 \times C_{12}^3 \times C_{12}^3 \times C_{12}^2$
- B $C_{12}^4 + C_8^3 + C_5^3 + C_2^2$
- C $C_{12}^4 \times 2 \times C_8^3 \times C_5^2$
- D $C_{12}^4 + 2 \times C_{12}^3 + C_{12}^2$
- E $C_{12}^4 \times C_8^3 \times C_5^3 \times C_2^2$

Comment: This problem seems to require previous knowledge of the combinatorial notation for combinations ($C_k^n = C(n,k) =$ “n choose k”) to be solved.

Possible answer: Among the 12 numbered wagons in the training, 4 must be painted red, 3 blue, 3 green, and 2 yellow. One can start by choosing which will be painted red. A combinatorial notation for this number (how many ways to choose 4 from 12 distinct objects) is C_{12}^4 . For each choice of which wagons will be painted red, one can then choose which of the remaining wagons will be painted blue, and so forth. This procedure is summarized as follows:

- choose 4 of the 12 wagons to be red - C_{12}^4
- then choose the 3 blue out of the remaining 8 wagons - C_8^3
- then choose the 3 green out of the remaining 4 wagons - C_5^3
- then “choose” the 2 yellow out of the 2 wagons - C_2^2

We get $C_{12}^4 \times C_8^3 \times C_5^3 \times C_2^2$ different possibilities for how to paint the wagons (Answer E).

Problem 11 - Motorbike Trip (Algebra)

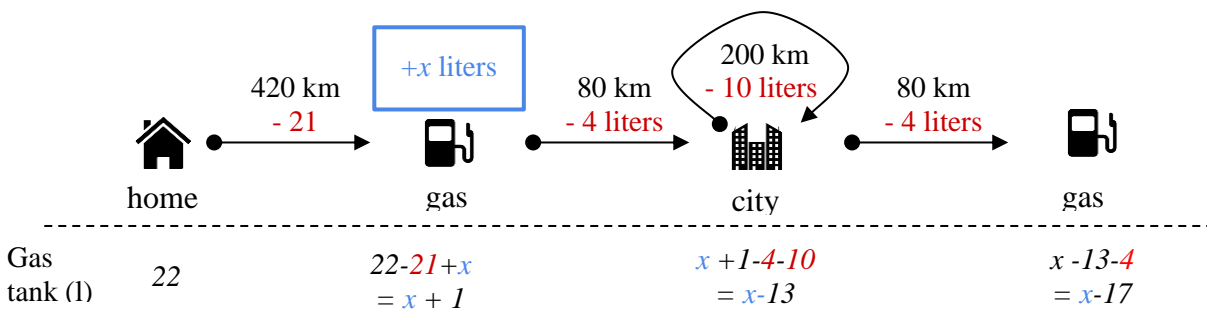
A motorcyclist plans to take a trip with a destination 500 km from his home. His motorcycle consumes 5 liters of gasoline for every 100 km traveled, and the motorcycle's tank has a capacity of 22 liters. On the map, he noted the last available refueling station on the journey, called Estrela, is 80 km from his destination. He intends to leave with a full motorcycle tank and plans to make only two stops for refueling, one on the way out and the other on the way back, both at the Estrela gas station. When refueling for the outward journey, he should also consider enough fuel to ride 200 km in the destination city.

The minimum amount of fuel, in liters, that this motorcyclist must refill at the Estrela gas station on the outward journey, which is sufficient to make the second refueling, is

- a) 13. b) 14. c) 17. d) 18. e) 21.

Comment: Solving this problem likely involves prior knowledge of proportions (how much fuel is needed for a given distance). As a long problem with many relevant numeric and contextual inputs, it seems challenging to keep track of all information while performing the necessary calculations.

Possible answer: Let x be the amount of gas filled at the first stop in the gas station. Reckoning that he starts with a full tank (22 liters) and calculating the amount of gas used in each part of the journey, we can figure he will have $x-17$ liters of gas in the tank on his second stop at the gas station.

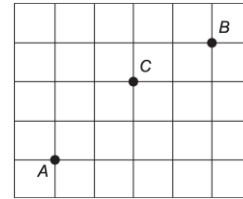


The problem asks us to find the minimum amount of gas to be filled in the first stop (x liters) that allows the driver to get to the second gas station stop. In this limit case, the amount of gas available in the tank when he makes the second stop must be zero. We thus get

$$x - 17 = 0 \rightarrow x = 17 \text{ (Answer C)}$$

Problem 12 - Paths on Map (Combinatorics)

Three friends, André, Bernardo, and Carlos, live in a closed condominium in a city. The grid represents the location of the parallel and perpendicular streets, delimiting blocks of the same size in this condominium. The houses of André, Bernardo, and Carlos, respectively, are located at points A, B, and C.



André wants to go from his house to Bernardo's house without going through Carlos' house, following along the streets of the condominium, always moving to the right (\rightarrow) or upwards (\uparrow), according to the diagram in the figure. The number of different paths that André can use to carry out the displacement under the proposed conditions is

- b) 4 b) 14 c) 17 d) 35 e) 48

Comment: This question has a small numerical answer and thus, in principle, can be solved by listing all cases. Keeping track of all cases seems to be a challenging task, though. Analytical solutions seem to also be challenging, requiring prior knowledge in combinatorics - more specifically, knowing how to code discrete map choices as combinations.

Possible answers:

Analytical solution using combinations:

“**All cases**”, from A to B you move 3 steps **up** and 4 steps **right**. This is equivalent to “choosing 3 steps among the 7” to be the “up” steps:

- $C(7,3) = \frac{7 \cdot 6 \cdot 5}{(3 \cdot 2 \cdot 1)} = 35$ ways

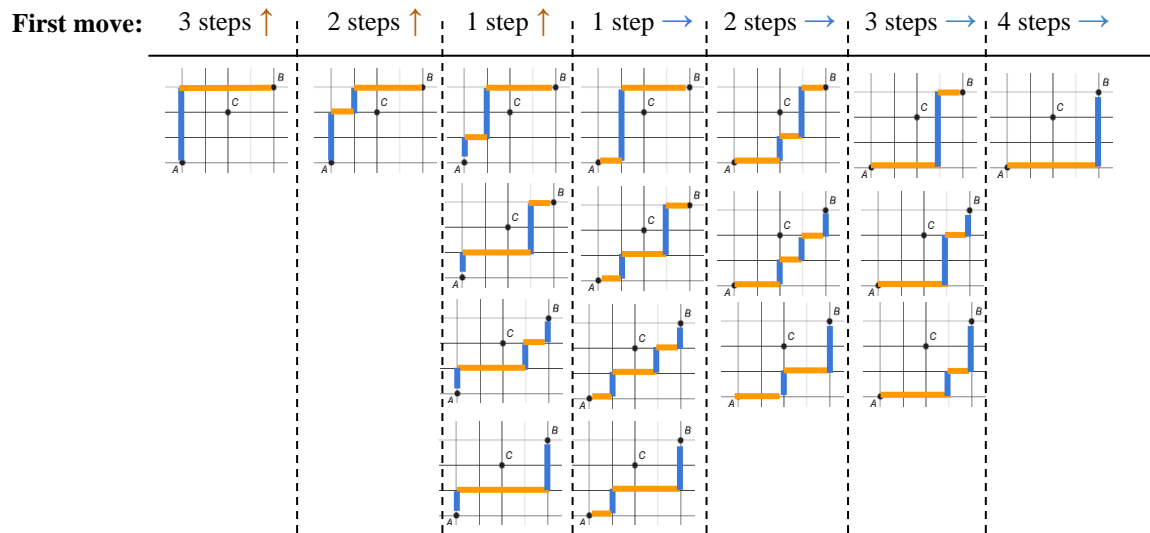
“**Bad cases**”, which are the paths $(A \rightarrow B) \times (C \rightarrow B)$.

- $C(4,2) \times C(3,1) = 6 \times 3 = 18$ ways

“**Good cases**” (*#all cases minus #bad cases*):

$35 - 18 = 17$ cases (Answer C).

Full list of cases:



Appendix C: Additional Tables

Table C.1: Regressions of private-public school achievement gaps (an indicator of accessibility, in the opposite direction) on a Combinatorics dummy and covariates using nearest neighbor matching samples.

	Model 1	Model 2	Model 3	Model 4
(Intercept)	0.116*** (0.013)	0.083 (0.052)	0.286 (0.151)	-0.014 (0.251)
combinatorics	-0.041* (0.018)	-0.041* (0.019)	-0.037 (0.019)	-0.038* (0.016)
year: 2010/2009		-0.020 (0.072)	-0.025 (0.071)	-0.067 (0.061)
year: 2011/2009		0.062 (0.062)	0.069 (0.062)	0.051 (0.052)
year: 2012/2009		0.057 (0.059)	0.046 (0.059)	0.038 (0.049)
year: 2013/2009		0.022 (0.062)	0.018 (0.062)	-0.020 (0.053)
year: 2014/2009		0.030 (0.072)	0.013 (0.072)	0.024 (0.060)
year: 2015/2009		0.019 (0.062)	0.006 (0.062)	0.011 (0.052)
year: 2016/2009		0.016 (0.059)	0.006 (0.059)	0.047 (0.050)
year: 2017/2009		0.044 (0.057)	0.029 (0.057)	0.049 (0.051)
year: 2018/2009		0.007 (0.072)	0.016 (0.071)	-0.022 (0.060)
year: 2019/2009		0.019 (0.062)	0.026 (0.062)	0.020 (0.052)
year: 2020/2009		0.044 (0.055)	0.040 (0.054)	0.034 (0.046)
CO_POSICAO			-0.001 (0.001)	
difficulty				0.873 (0.750)
difficulty2				-0.950 (0.563)
R-squared	0.087	0.169	0.208	0.455
N	56	56	56	56

Significance: *** = $p < 0.001$; ** = $p < 0.01$; * = $p < 0.05$

Table C.2: Regressions of white-black student achievement gaps (an indicator of accessibility, in the opposite direction) on a Combinatorics dummy and covariates using nearest neighbor matching samples.

	Model 1	Model 2	Model 3	Model 4
(Intercept)	0.055*** (0.006)	0.036 (0.024)	0.095 (0.069)	0.116 (0.100)
combinatorics	-0.022* (0.008)	-0.022* (0.009)	-0.020* (0.009)	-0.021** (0.006)
year: 2010/2009		0.000 (0.033)	-0.001 (0.033)	-0.022 (0.024)
year: 2011/2009		0.024 (0.028)	0.026 (0.029)	0.017 (0.021)
year: 2012/2009		0.026 (0.027)	0.023 (0.027)	0.016 (0.020)
year: 2013/2009		0.022 (0.028)	0.020 (0.028)	0.001 (0.021)
year: 2014/2009		0.016 (0.033)	0.011 (0.033)	0.013 (0.024)
year: 2015/2009		0.008 (0.028)	0.004 (0.029)	0.005 (0.021)
year: 2016/2009		0.004 (0.027)	0.001 (0.027)	0.017 (0.020)
year: 2017/2009		0.026 (0.026)	0.022 (0.026)	0.018 (0.020)
year: 2018/2009		0.004 (0.033)	0.007 (0.033)	-0.009 (0.024)
year: 2019/2009		0.015 (0.028)	0.017 (0.028)	0.015 (0.021)
year: 2020/2009		0.028 (0.025)	0.027 (0.025)	0.023 (0.018)
CO_POSICAO			-0.000 (0.000)	
difficulty				0.078 (0.299)
difficulty2				-0.226 (0.224)
R-squared	0.112	0.209	0.224	0.601
N	56	56	56	56
Significance: *** = p < 0.001; ** = p < 0.01; * = p < 0.05				

Table C.3: Regression models including observable characteristics (*length* and *notation*) for the sample of non-Combinatorics problems only. All regressions include year fixed effects, even though their coefficients are not reported.

	<i>Difficulty</i>	<i>School gap</i>	<i>Race gap</i>
<i>(Intercept)</i>	0.408*** (0.062)	0.116* (0.053)	0.116*** (0.024)
<i>item position</i>	0.005 (0.003)	-0.000 (0.002)	-0.000 (0.001)
<i>length</i>	0.002*** (0.000)	-0.001*** (0.000)	0.000** (0.000)
<i>notation</i>	0.065*** (0.013)	-0.020** (0.008)	-0.011** (0.004)
<i>difficulty</i>		0.882*** (0.149)	0.262*** (0.068)
<i>difficulty sq.</i>		-0.958*** (0.116)	-0.391*** (0.053)
<i>year fixed effects</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
<i>R-squared</i>	<i>0.151</i>	<i>0.445</i>	<i>0.568</i>
<i>N (items)</i>	<i>600</i>	<i>600</i>	<i>600</i>

Significance: *** = $p < 0.001$; ** = $p < 0.01$; * = $p < 0.05$. Standard deviations are shown in parenthesis, below each coefficient.

Table C.4. Regression models including observable characteristics (*length* and *notation*) for the sample of Combinatorics problems only. All regressions include year fixed effects, even though their coefficients are not reported.

	<u><i>Difficulty</i></u>	<u><i>School gap</i></u>	<u><i>Race gap</i></u>
<i>(Intercept)</i>	-0.062 (0.485)	0.094 (0.439)	0.159 (0.187)
<i>item position</i>	0.005 (0.003)	-0.000 (0.002)	-0.000 (0.001)
<i>length</i>	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)
<i>notation</i>	0.063 (0.063)	0.044 (0.035)	0.014** (0.015)
<i>difficulty</i>		0.348 (1.151)	-0.161 (0.490)
<i>difficulty sq.</i>		-0.581 (0.863)	-0.033 (0.367)
<i>year fixed effects</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
<i>R-squared</i>	<i>0.328</i>	<i>0.673</i>	<i>0.685</i>
<i>N (items)</i>	<i>28</i>	<i>28</i>	<i>28</i>

Significance: *** = $p < 0.001$; ** = $p < 0.01$; * = $p < 0.05$. Standard deviations are shown in parenthesis, below each coefficient.