



Universidade de Brasília

Instituto de Ciências Exatas
Departamento de Ciência da Computação

ChatGPT como um tutor para o ensino de Algoritmos e Programação de Computadores: Um relato de experiência

Camila Frealdo Fraga
Luiza de Araújo Nunes Gomes

Monografia apresentada como requisito parcial
para conclusão do Bacharelado em Ciência da Computação

Orientadora
Prof.a Dr.a Maristela Terto de Holanda

Brasília
2025



Universidade de Brasília

Instituto de Ciências Exatas
Departamento de Ciência da Computação

ChatGPT como um tutor para o ensino de Algoritmos e Programação de Computadores: Um relato de experiência

Camila Frealdo Fraga
Luiza de Araújo Nunes Gomes

Monografia apresentada como requisito parcial
para conclusão do Bacharelado em Ciência da Computação

Prof.a Dr.a Maristela Terto de Holanda (Orientadora)
CIC/UnB

Prof. Dr. Edison Ishikawa Prof. Dr. Jonathan Rosa Moreira
Universidade de Brasília (UnB) Universidade de Brasília (UnB)

Prof. Dr. Marcelo Grandi Mandelli
Coordenador do Bacharelado em Ciência da Computação

Brasília, 13 de fevereiro de 2025

Dedicatória

Camila: Dedico este trabalho a todos que me acompanharam e contribuíram para minha jornada ao longo do curso de Ciência da Computação.

Luiza: Eu dedico este trabalho à minha família, aos meus amigos e ao meu amado cachorro Sasuke, que me acompanhou ao longo desta longa trajetória do curso de Ciência da Computação.

Agradecimentos

Camila: Este trabalho e minha trajetória no curso de Ciência da Computação foram jornadas de autoconhecimento e muitos desafios, e só foram possíveis graças ao apoio de pessoas que sempre estiveram ao meu lado. Cada dificuldade enfrentada no curso e vencida tornou essa conquista ainda mais importante. Agradeço aos meus pais pelo incentivo e por confiarem no meu potencial, por me darem a oportunidade de chegar até aqui, e à minha irmã, pelo apoio e por estar sempre presente nos momentos mais difíceis. Ao meu namorado, que sempre acreditou em mim, mesmo quando eu mesma duvidava, que esteve ao meu lado nos momentos mais difíceis, e que sempre me incentivou a persistir e confiar mais em mim. Agradeço aos professores por todos os ensinamentos, especialmente à professora Maristela por dedicar seu tempo, paciência e suporte, que foram essenciais para a realização deste trabalho. Agradeço também ao professor Edison pela oportunidade de realizar os experimentos na disciplina de APC.

Luiza: O final do meu curso foi um momento tão intenso em minha vida que fica difícil agradecer a todas as pessoas que me ajudaram a chegar até aqui. Início com um agradecimento mais que especial ao Sasuke, à minha família e às minhas amigas e amigos. Afinal, cada um me ajudou a construir, bloquinho por bloquinho, desta graduação, deste trabalho de conclusão de curso e da minha vida universitária. Agradeço também à professora Maristela por dedicar seu tempo e disponibilidade para nos orientar ao longo desta trajetória no trabalho de conclusão de curso e por nos proporcionar a oportunidade de desenvolver um estudo tão relevante para a comunidade acadêmica. Por fim, agradeço ao professor Edison Ishikawa por disponibilizar um espaço em sua disciplina para a realização dos experimentos deste estudo. Portanto, sou grata a mim mesma por persistir nesse caminho tão complicado e a todos que contribuíram para este momento, pois, de alguma forma, cada um me ajudou, à sua maneira, a alcançar essa conquista.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES), por meio do Acesso ao Portal de Periódicos.

Resumo

ChatGPT se tornou uma das ferramentas de inteligência artificial mais utilizadas entre estudantes no setor da programação. Destaca-se no auxílio à aprendizagem, desde cálculo até programação. Este artigo apresenta um relato de experiência de um experimento controlado conduzido na primeira disciplina de programação do primeiro semestre do curso de Licenciatura em Computação na Universidade de Brasília (UnB). O experimento teve como objetivo avaliar a performance e percepção dos alunos quanto ao uso do ChatGPT em sala de aula, e sua capacidade de assistência em comparação à tutoria tradicional, diferenciando-se de trabalhos anteriores ao realizar essa análise em uma turma de introdução à programação. Com base nesse contexto, este estudo busca responder às seguintes questões: **QP1** - Quais são as diferenças no impacto da tutoria baseada no ChatGPT e da tutoria tradicional no aprendizado e desempenho dos estudantes em atividades de programação?; **QP2** - Qual a percepção dos estudantes sobre o uso do ChatGPT nas atividades em sala de aula?; e **QP3** - O uso do ChatGPT pode ter impacto na nota dos alunos na avaliação da disciplina?

O experimento contou com a participação de 41 estudantes e ocorreu ao longo de três semanas consecutivas e três dias de atividade com duração de 80 minutos cada, seguidos de um formulário pós-atividade que foi conduzido para coletar a percepção dos estudantes sobre o dia de atividades. As respostas dos exercícios submetidos pelos estudantes indicaram que os alunos que utilizaram o ChatGPT conseguiram submeter mais soluções de exercícios em menos tempo do que o grupo controle. No entanto, as soluções submetidas pelo grupo do ChatGPT não estavam necessariamente corretas. Os resultados dos formulários indicaram que os estudantes acreditam que o ChatGPT pode ser uma ferramenta útil em sala de aula. Ainda assim, também expressaram uma preferência pela tutoria tradicional em comparação à assistência do ChatGPT.

Palavras-chave: ChatGPT, Tutoria, Ensino de Programação, Chatbot de IA

Abstract

ChatGPT has become one of the most widely used artificial intelligence tools among students in the programming sector. It assists with learning in areas ranging from calculus to programming. This paper presents an experience report for a controlled experiment conducted in the first programming course of the Teaching Degree in Computer Science at the University of Brasília (UnB). The experiment aimed to evaluate students' performance and perception of the use of ChatGPT in class and its capability of assisting in comparison to traditional tutoring, distinguishing itself from previous studies by conducting this analysis in an introductory programming class. Given this context, this study seeks to answer the following research questions: **RQ1** - What are the differences in the impact of ChatGPT-based tutoring and traditional tutoring on students' learning and performance in programming activities?; **RQ2** - What are the students' perception of using ChatGPT in classroom activities?; and **RQ3** - Can the use of ChatGPT impact students' grades in the course evaluation?

The experience involved the participation of 41 students. It consisted of three days of activities throughout three consecutive weeks, each activity lasting 80 minutes long, followed by post-activity surveys conducted to collect the students' perceptions of the day's activity. The students' submitted exercise answers indicated that those students using ChatGPT were able to submit more exercise solutions in less time than the students in the control group. However, the solutions submitted by the ChatGPT group were not necessarily more accurate. Both groups achieved a similar accuracy rate for the exercises they submitted. The survey results indicated that students believed that ChatGPT could be useful in class. Still, they expressed a preference for traditional tutoring over assistance from ChatGPT.

Keywords: ChatGPT, Tutoring, Programming Education, AI-Chatbot

Sumário

1	Introdução	1
2	Referencial Teórico	3
3	Trabalhos Relacionados	5
4	Disciplina Algoritmos e Programação de Computadores	8
5	Metodologia dos Experimentos	10
6	Perfil da Turma	13
6.1	Divisão dos Grupos	13
7	ChatGPT x Tutoria Tradicional	16
7.1	Participação e Desempenho dos Alunos nos Experimentos	16
7.2	Análise do Dia 1	17
7.3	Análise do Dia 2	18
7.4	Análise do Dia 3	19
7.5	Análise Geral	21
8	Percepção dos estudantes sobre o uso do ChatGPT	24
8.1	Percepções do Dia 1	24
8.2	Percepções do Dia 2	26
8.3	Percepções do Dia 3	29
8.4	Percepções Gerais	31
9	O uso do ChatGPT pode ter impacto na nota dos alunos na avaliação da disciplina?	34
10	Lições Aprendidas	36
11	Limitações	38

12 Conclusão	40
Referências	42
Apêndice	44
A Listas de Exercícios	45
A.1 Lista de Exercícios - Dia 1	45
A.1.1 Exercício 1 (<i>Equivalente ao Beecrowd 1038</i>)	45
A.1.2 Exercício 2 (<i>Equivalente ao Beecrowd 1051</i>)	46
A.1.3 Exercício 3 (<i>Equivalente ao Beecrowd 1117</i>)	46
A.2 Lista de Exercícios - Dia 2	47
A.2.1 Exercício 1 (<i>Equivalente ao Beecrowd 2717</i>)	47
A.2.2 Exercício 2 (<i>Equivalente ao Beecrowd 1150</i>)	48
A.2.3 Exercício 3 (<i>Equivalente ao Beecrowd 1146</i>)	48
A.3 Lista de Exercícios - Dia 3	49
A.3.1 Exercício 1 (<i>Equivalente ao Beecrowd 1151</i>)	49
A.3.2 Exercício 2 (<i>Equivalente ao Beecrowd 3358</i>)	49
A.3.3 Exercício 3 (<i>Equivalente ao Beecrowd 2137</i>)	50
B Listas de <i>Prompts</i>	51
B.1 Para dúvidas com o Exercício	51
B.1.1 <i>Prompt 1</i>	51
B.1.2 <i>Prompt 2</i>	51
B.1.3 <i>Prompt 3</i>	51
B.2 Para depuração de código	51
B.2.1 <i>Prompt 4</i>	51
B.3 Ajuda com ideias e melhorias de código	52
B.3.1 <i>Prompt 5</i>	52
B.3.2 <i>Prompt 6</i>	52

Lista de Figuras

6.1	Dados referentes ao perfil dos alunos coletados durante a pesquisa.	14
7.1	Percentual de Erros por Tipo de Erro - Comparação entre os grupos Controle e Experimental no Dia 1	18
7.2	Percentual de Erros por Tipo de Erro - Comparação entre os grupos Controle e Experimental no Dia 2	19
7.3	Percentual de Erros por Tipo de Erro - Comparação entre os grupos Controle e Experimental no Dia 3	20
7.4	Comparação de dificuldade e dúvidas - Dia 1.	22
7.5	Comparação de dificuldade e dúvidas - Dia 2.	23
7.6	Comparação de dificuldade e dúvidas - Dia 3.	23
8.1	Sentimento geral sobre o ChatGPT.	27
8.2	Sentimento sobre o uso do ChatGPT na atividade em sala de aula.	28
8.3	Nuvem de palavras dos comentários sobre o uso do ChatGPT em sala de aula.	32
9.1	Distribuição das Menções da Prova	35

Lista de Tabelas

5.1	Descrição dos tipos de erros.	12
7.1	Distribuição semanal dos conteúdos ministrados e quantidade de alunos presentes nos grupos controle e experimental	16
7.2	Comparativo entre os grupos controle e experimental quanto ao término e acerto dos exercícios aplicados na Dia 1	17
7.3	Comparativo entre os grupos controle e experimental quanto ao término e acerto dos exercícios aplicados na Dia 2	18
7.4	Comparativo entre os grupos controle e experimental quanto ao término e acerto dos exercícios aplicados na Dia 3	20
7.5	Relação entre o número de resoluções entregues e o total de questões que deveriam ser resolvidas em cada dia de experimento	21
7.6	Proporção entre o número de resoluções corretas em relação as questões enviadas em cada dia do experimento	21
8.1	Resultados referentes às questões de sentimento geral sobre o ChatGPT no formulário pós-atividade.	25
8.2	Resultados referentes às questões de sentimento sobre o uso do ChatGPT no formulário pós-atividade.	26
8.3	Perguntas abertas sobre percepção dos estudantes.	33
9.1	Estatísticas Descritivas dos Resultados da Prova	35
A.1	Relação de código, especificação e preço do produto	45
A.2	Casos de Teste do Exercício A.1.1	45
A.3	Relação entre faixa de renda e taxa de imposto de renda	46
A.4	Casos de Teste do Exercício A.1.2	46
A.5	Casos de Teste do Exercício A.1.3	47
A.6	Casos de Teste do Exercício A.2.1	47
A.7	Casos de Teste do Exercício A.2.2	48
A.8	Casos de Teste do Exercício A.2.3	49

A.9 Casos de Teste do Exercício A.3.1	49
A.10 Casos de Teste do Exercício A.3.2	50
A.11 Casos de Teste do Exercício A.3.3.	50

Capítulo 1

Introdução

Os estudos na área de Inteligência Artificial (IA) têm feito progressos relevantes nos últimos anos, tornando possíveis muitas oportunidades e transformando diversos campos da vida cotidiana, incluindo a educação. O ChatGPT vem surgindo como uma ferramenta potencial para otimizar algumas tarefas, representando o estado da arte da IA e impulsionando a reforma da educação [1].

Lançado pela OpenAI em Novembro de 2022, o ChatGPT é um *chatbot* baseado em um modelo de LLM (*Large Language Model*) chamado GPT (*Chat Generative Pre-trained Transformer*). Este possui uma capacidade conversacional fluida, que com auxílio do grande volume de dados utilizados como base de conhecimento, transformou o software em uma das aplicações mais populares em poucos anos pela sua capacidade de conversação elevada e o grande volume de informação.

Além de suas características técnicas, o ChatGPT ganhou mais usuários pela facilidade do uso de sua interface, pois estes podem conversar em linguagem natural com o *chatbot* e receber diversos tipos de respostas, que podem ser um simples cálculo à uma geração de código de programação. Ou seja, o ChatGPT é uma ferramenta bastante versátil e simples de utilizar, sendo cada vez mais empregada em diferentes segmentos da sociedade, como, por exemplo, na educação.

No contexto educacional, o ChatGPT apresenta um grande potencial para auxiliar professores e alunos em atividades de ensino-aprendizagem, como por exemplo a resolução de problemas de programação, elaboração de explicações detalhadas e geração de códigos em diferentes linguagens de programação. Além disso, o seu uso tem o potencial benefício de fornecer *feedback* instantâneo para tarefas de programação, o que permite que os alunos identifiquem e corrijam erros de maneira rápida, o que facilita o processo de aprendizagem, se utilizado com cautela [2] [3].

Dessa forma, enquanto, por um lado o ChatGPT vem se destacando como uma ferramenta educacional, funcionando como uma espécie de monitor virtual sempre disponível,

com uma vasta quantidade de informação e uma linguagem acessível. Por outro, por ser capaz de gerar resoluções para atividades educacionais com rapidez e facilidade, pode representar um risco ao processo de aprendizado, reduzindo a capacidade dos alunos resolverem os exercícios de forma independente.

Portanto, com o aumento do uso de ferramentas baseadas em IA, como o ChatGPT, no ensino superior e em cursos de programação é importante a análise de seu impacto. Apesar do interesse acadêmico no uso da ferramenta, ainda há poucos estudos empíricos sobre seu impacto real no desempenho prático de programação, especialmente no contexto universitário brasileiro. O ensino de programação apresenta um desafio para muitos alunos e o uso do ChatGPT pode ajudar ou não a superar esses desafios. Com isso é importante a discussão sobre a integração dessas tecnologias de forma ética e responsável.

Este estudo apresenta um relato de experiência sobre o uso do ChatGPT em sala de aula na disciplina de Algoritmos e Programação de Computadores (APC), oferecida no curso de Licenciatura em Computação da Universidade de Brasília (UnB). O objetivo deste relato de experiência é avaliar o impacto do uso do ChatGPT em sala de aula no desempenho dos estudantes em comparação com a tutoria tradicional, bem como compreender a percepção dos participantes sobre o uso da ferramenta de IA em atividades de programação em sala de aula. Este trabalho pode ajudar professores e instituições a entender melhor o impacto dessa ferramenta no desempenho dos estudantes além de fornecer insights sobre as percepções dos alunos sobre o uso da ferramenta em ambiente acadêmico.

O trabalho foi dividido nos seguintes capítulos: o Capítulo 2 apresenta o referencial teórico; o Capítulo 3 apresenta os trabalhos relacionados; o Capítulo 4 apresenta a disciplina Algoritmos e Programação de Computadores; o Capítulo 5 apresenta a metodologia utilizada; o Capítulo 6 apresenta o perfil da turma; os Capítulos 7, 8 e 9 apresentam os resultados obtidos e discussões; o Capítulo 10 apresenta as lições aprendidas; o Capítulo 11 apresenta as limitações; e por fim, o Capítulo 12 apresenta as conclusões e sugestões para trabalhos futuros.

Capítulo 2

Referencial Teórico

A inteligência artificial tem sido cada vez mais comum no contexto educacional, sendo que pode tanto auxiliar quanto representar desafios para alunos e professores. Um lugar de destaque está no ensino superior. Ferramentas de *chatbot* como no estilo ChatGPT, Gemini e outros modelos de linguagem têm sido utilizadas como maneira de melhorar o engajamento e o desempenho dos alunos [4]. Outra maneira que essas ferramentas vêm sendo utilizadas é na personalização do aprendizado, tornando o aprendizado mais acessível aos estudantes [5] [6] [7]. No período pós-pandemia, a IA influenciou a importância do desenvolvimento de novos métodos de avaliação, deixando evidente a necessidade de mudança nas práticas acadêmicas [8] [9] [10] [1].

Desde o seu lançamento, o ChatGPT vem sendo amplamente estudado por seu potencial de auxiliar em diversas áreas e tarefas educacionais como matemática, física, programação e etc. Algumas tarefas na área de programação incluem a explicação de códigos, correção de exercícios de programação, geração de *feedback* personalizado [2] [11]. Além de otimização e sugestões de refatoração de código [12]. Em alguns cursos específicos como o de programação funcional em Scala, o ChatGPT mostrou um desempenho satisfatório, mesmo apresentando algumas limitações em casos de tarefas mais complexas e também mostrou potencial como ferramenta para automatizar e auxiliar na revisão de código [11]. Professores podem usar o ChatGPT para criar rubricas de avaliação que expliquem de forma clara o que é necessário ser feito pelos alunos [10]. Além disso, o seu potencial criativo foi avaliado, como no caso do design de jogos, o que mostra que sua aplicação pode ir além de apenas tarefas técnicas [13].

Alguns estudos analisaram como o ChatGPT pode impactar no ensino e aprendizagem na área de programação. A ferramenta tem sido empregada tanto para auxiliar na compreensão de conceitos avançados sobre algoritmos num âmbito geral quanto para facilitar o aprendizado em linguagens de programação como PHP [14] [15]. Entretanto, essas e outras pesquisas também destacam as limitações do *chatbot*, como as “alucinações” nas

quais a ferramenta pode apresentar falsas correções apresentando erros inexistentes [2] ou gerar respostas erradas e inventar artigos [10]. Além disso, outros estudos analisaram como o ChatGPT pode influenciar o comportamento e o desempenho dos alunos, mostrando como ele pode ser útil para explicações e correções, mas também pode afetar o aprendizado ao introduzir certa dependência [16].

Apesar do uso da IA generativa apresentar benefícios no ensino superior, ele também apresenta limitações e preocupações. Pode-se citar algumas, tais como: o risco de plágio, a dependência excessiva de ferramentas e as preocupações relacionadas à formação do pensamento crítico dos estudantes [3] [12] [17]. Dessa forma, fica evidente a importância do desenvolvimento de habilidades críticas e da literacia em IA nos estudantes e professores para o uso de maneira responsável dessas tecnologias [18]. Verifica-se que há variação na percepção de professores e alunos sobre o uso de IA no ambiente educacional, porém há consenso acerca da importância da integração dessas ferramentas, de forma responsável, para usufruir dos potenciais benefícios para o aprendizado [9].

Alguns estudos como [19] [3] apontam que, para que ferramentas como essas sejam utilizadas em sala de aula, é essencial que sejam utilizadas com precaução, de maneira controlada. Assim, os alunos devem ser adequadamente orientados sobre a forma correta de utilizá-las, evitando que valores básicos do ensino superior sejam violados. Portanto, é de suma importância que, para aproveitar os potenciais benefícios dessas ferramentas, é preciso esforço conjunto entre as instituições, professores e alunos na exploração do uso ético dessas tecnologias [20] [10].

Capítulo 3

Trabalhos Relacionados

O potencial do ChatGPT vem sendo explorado em diversos estudos com o intuito de avaliar se a capacidade e a qualidade do suporte fornecido pelo *chatbot* são eficientes. Um dos pontos analisados foi a capacidade do ChatGPT de responder de maneira similar à forma como um humano responderia, considerando que a LLM é capaz de responder perguntas, completar textos e realizar outras tarefas baseadas em linguagem natural. Ou seja, estuda-se se a habilidade do ChatGPT de processar e gerar textos semelhantes aos humanos é confiável, pois nota-se que o conhecimento da ferramenta não é equivalente ao entendimento humano. Suas limitações reforçam a necessidade de que os usuários avaliem criticamente as respostas fornecidas.

Dessa forma, no estudo realizado em [21], observou-se que o *chatbot* foi solicitado a responder às perguntas abertas e fechadas (múltipla escolha) do Exame Nacional de Desempenho dos Estudantes (ENADE) de 2021 para os alunos de Ciência da Computação. O intuito do estudo era avaliar a qualidade das respostas fornecidas pelo ChatGPT; também, como sairia caso errasse as respostas e fosse confrontado com as resoluções; se seria capaz de ser aprovado no exame; e como seu desempenho se compararia ao dos estudantes que realizaram o mesmo exame. De modo geral, o ChatGPT obteve um bom desempenho nos dois tipos de perguntas do exame. No entanto, apresentou maior dificuldade em responder questões que possuíam respostas em várias etapas. Quando confrontado com respostas corretas, o *chatbot* foi capaz de alterar algumas respostas, mas nem sempre o fez. Em alguns casos, o *chatbot* manteve as respostas anteriores, mesmo com a oportunidade de reavaliá-las. Ainda assim, o *chatbot* obteve uma nota superior à média dos estudantes que realizaram o exame.

Outro ponto também bastante analisado na literatura foi a efetividade do uso de LLMs, como o ChatGPT como ferramentas de aprendizado para programação. Vários estudos foram conduzidos [22] [23] [15], em que um conjunto de alunos foi separado em dois grupos: um controle, em que não utilizaria o ChatGPT e outro experimental, que o

utilizaria. Esses estudos procuraram entender a percepção dos alunos quanto ao uso do ChatGPT, a performance entre os grupos e a efetividade da ferramenta como suporte à aprendizagem de programação disciplinas introdutórias de programação.

No estudo realizado em [22], os alunos de uma turma de uma disciplina de introdução à programação, em Python, foram separados em dois grupos igualitário, considerando o pré-conhecimento de Python de cada aluno. Os alunos que utilizaram o ChatGPT participaram de uma aula expositiva sobre o uso do *chatbot* e uma breve discussão sobre suas limitações da ferramenta e questões de plágio. Em seguida, os dois grupos realizaram uma mesma atividade, sendo solicitado que os alunos submetessem uma transcrição de suas interações com o ChatGPT. Após o término da atividade, todos os alunos responderam um formulário avaliando o uso do ChatGPT e sobre a experiência de participar da atividade. Após isso, os dados foram analisados, e a maioria dos alunos classificou o ChatGPT como uma ferramenta útil, relatando que adquiriram conhecimento a partir da interação. Além disso, ao corrigir as atividades, foi observado que as respostas dos alunos que utilizaram o ChatGPT eram mais detalhadas e continham menos erros gramaticais.

No estudo de [23], em uma atividade proposta em uma disciplina também de introdução à programação, mas em Java, os alunos puderam escolher a qual dos dois grupos à seguir gostariam de pertencer: um com acesso a recursos online de ensino (como Google e apresentações do curso), mas sem uso do ChatGPT ou outra IA, e outro com acesso tanto aos recursos online quanto ao ChatGPT. Durante a atividade, os alunos deveriam gravar as suas telas, e ao final, os alunos responderam a uma pesquisa sobre detalhes demográficos e habilidades relacionadas à computação. Também, foram coletados os *feedbacks* sobre a percepção e as opiniões dos alunos quanto à influência do ChatGPT na disciplina. Em síntese, os alunos reconheceram alguns benefícios no uso do ChatGPT, como para depuração de código ou geração de ideias. No entanto, relataram preocupações éticas, especialmente em relação à dependência excessiva e ao potencial impacto no próprio aprendizado.

Os autores de [23] observaram que os professores devem adotar precaução ao integrar o ChatGPT em disciplinas. A recomendação principal foi utilizar o ChatGPT como uma ferramenta suplementar, e não como substituto dos recursos existentes. Em outro estudo [24], essa recomendação foi colocada em prática em uma disciplina de "Estruturas de Dados", na qual os alunos foram incentivados a usar o ChatGPT como suporte ao longo do semestre. Os autores geraram e forneceram *prompts*, para aprendizado de árvores binárias e revisados pelo professor da disciplinas. Ao final do semestre, os alunos avaliaram o ChatGPT quanto à efetividade da ferramenta nos estudos. Apesar de relatarem casos de códigos e explicações erradas, os autores concluíram que a ferramenta supera suas falhas, é benéfica como ferramenta de apoio, desde que seja utilizada com devida cautela.

Quanto à performance e à qualidade do aprendizado de alunos iniciantes em programação utilizando o ChatGPT, destaca-se o estudo apresentado em [15]. Neste trabalho, além de realizar uma avaliação da performance dos alunos, houve uma inversão dos grupos em dois questionários diferentes. Assim, os dois grupos, divididos de forma randômica, responderam um questionário utilizando o ChatGPT e outro manualmente, garantindo que todos tivessem a oportunidade de utilizar a ferramenta. Os resultados indicaram que os alunos que utilizaram o ChatGPT pontuaram significativamente mais alto em questões de codificação em comparação aos alunos que responderam manualmente. Os autores concluíram que o ChatGPT pode melhorar o desempenho dependendo dos estilos de questões aplicadas, destacando o benefício em atividades de codificação.

Assim, observa-se que há estudos que avaliam a capacidade e qualidade do suporte fornecido pelo ChatGPT; que analisam a efetividade do uso de LLMs no aprendizado de programação; e ainda, outros que avaliam a performance e a qualidade do aprendizado de alunos iniciantes em programação que utilizam o *chatbot*. Contudo, nenhum desses avaliam o impacto do uso do ChatGPT como um tutor em comparação à tutoria tradicional (monitores, tutores e professores em sala). Além disso, não há estudos que investiguem o impacto da ferramenta em avaliações formais de uma disciplina, como foi realizado no experimento relatado neste artigo. Neste trabalho, além dos experimentos em sala de aula, analisou-se o resultado obtido pelos alunos em uma avaliação na qual os alunos não tiveram auxílio externo (tutoria), para que fossem avaliados com base no conhecimento adquirido ao utilizar o *chatbot* em sala de aula.

Capítulo 4

Disciplina Algoritmos e Programação de Computadores

A primeira disciplina de programação do curso de Licenciatura em Computação chama-se Algoritmos e Programação de Computadores; esta disciplina tem como objetivos apresentar aos estudantes os princípios fundamentais de construção de programas; ensinar noções de abstração computacional de forma que os estudantes possam modelar os problemas propostos de forma clara e eficiente, utilizando variáveis, funções e estruturas de controle; e desenvolver a habilidade dos alunos de resolver problemas, aplicando os conceitos de programação. A linguagem de programação utilizada no segundo semestre de 2024 foi a linguagem Python.

A disciplina é ministrada presencialmente na Universidade de Brasília (UnB) e é conduzida por meio de atividades presenciais em sala de aula e nos Laboratórios de Informática (LINF), que dispõem de computadores com acesso à internet para todos os estudantes matriculados na disciplina. O conteúdo da disciplina é dividido em dez tópicos: Algoritmos e subalgoritmos; Variáveis, valores e tipos; Estruturas sequenciais; Estruturas condicionais; Estruturas de repetição; Strings; Listas e Tuplas; Funções; Dicionários; e Recursividade.

As atividades presenciais são compostas por aulas teóricas realizadas em sala de aula às terças-feiras e aulas práticas no Laboratório de Informática às quartas e quintas-feiras, dedicadas à resolução de exercícios, fixação de conteúdo e retirada de dúvidas com professores, monitores e tutores. Cada aula tem duração de 100 minutos, além das provas presenciais programadas com conteúdos predefinidos. A comunicação entre o docente, os estudantes, os monitores e os tutores ocorre preferencialmente por meio das ferramentas institucionais. Todo o material didático da disciplina, tais como leituras obrigatórias, videoaulas, listas de exercícios, avaliações e material complementar, está armazenado na plataforma educacional Aprender3, uma versão customizada do Moodle.

A avaliação da aprendizagem é feita por meio de duas provas presenciais, questionários

e um projeto. Cada questionário aborda um dos dez tópicos apresentados anteriormente e pode ser realizado com consulta. Para as provas, é permitida a consulta aos slides da disciplina disponíveis na plataforma Aprender3. O conteúdo da primeira prova aborda os tópicos Algoritmos e subalgoritmos; Variáveis, valores e tipos; Estruturas sequenciais; Estruturas condicionais; e Estruturas de repetição. Já o conteúdo da segunda prova e do projeto aborda todo o conteúdo programático da disciplina.

Capítulo 5

Metodologia dos Experimentos

Este trabalho visa avaliar o impacto do ChatGPT no aprendizado e desempenho dos alunos da primeira disciplina de programação em uma turma do curso de Licenciatura em Computação, bem como analisar a percepção dos estudantes quanto ao uso do assistente virtual em sala de aula. Para isso, foi adotado uma abordagem quantitativa dos dados obtidos durante o experimento. Assim, este estudo busca responder às seguintes questões de pesquisa (QP):

- **QP1:** Quais são as diferenças no impacto da tutoria baseada no ChatGPT e da tutoria tradicional no aprendizado e desempenho dos estudantes em atividades de programação?
- **QP2:** Qual a percepção dos estudantes sobre o uso do ChatGPT nas atividades em sala de aula?
- **QP3:** O uso do ChatGPT pode ter impacto na nota dos alunos na avaliação da disciplina?

Para responder às questões de pesquisa, inicialmente os estudantes preencheram um formulário com perguntas de respostas simples (Sim ou Não) e escala *Likert* de 1 a 5 elaborado na ferramenta Google Forms sobre familiaridade com o uso de ferramentas de inteligência artificial e noções de programação. Com base em suas respostas, os alunos foram divididos em dois grupos de forma equiparada, para evitar uma disparidade na avaliação final.

Dessa forma, os grupos foram divididos fisicamente em dois laboratórios diferentes, em um laboratório, o grupo experimental poderia utilizar como auxílio para a resolução dos exercícios propostos somente o ChatGPT como ferramenta de inteligência artificial, como também, poderiam realizar pesquisas online, utilizar os materiais disponíveis no Aprender3, e conversar com os colegas de sala. O grupo controle, poderia acessar tudo que

o grupo experimental acessou, com exceção do ChatGPT e outras inteligências artificiais, mas teria acesso aos monitores, tutores e docente durante a aula para sanar dúvidas.

Após a separação física dos grupos, os alunos do grupo experimental foram submetidos a uma aula expositiva (apenas no primeiro dia do experimento) com a apresentação de dicas de uso do ChatGPT e *prompts* de comando que poderiam ser utilizados durante as atividades. Ao todo foram disponibilizados seis *prompts* criados pelos autores deste artigo por meio da técnica “*metaprompting*”[25]: três para sanar dúvidas sobre o enunciado dos exercícios, um para depuração de código e dois para otimização de código. Os materiais ficaram disponíveis aos alunos durante todas as aulas. Os prompts gerados estão disponíveis no Apêndice B.3.2.

O experimento proposto foi realizado durante três semanas no mês de dezembro de 2024, nas aulas práticas de quartas-feiras no LINF na Universidade de Brasília. Ao todo, foram realizados três experimentos. Para cada dia, foi preparada uma atividade sobre o assunto ensinado na aula teórica do dia anterior com três exercícios de dificuldade fácil à média retirados do site *Beecrowd*¹, uma plataforma com diversas funcionalidades, sendo uma das principais a disponibilização de problemas de programação. Esses exercícios estão disponíveis no Apêndice A.3.3.

Os estudantes dos dois grupos foram orientados a concluírem a atividade em um tempo máximo de 80 minutos. Após o fim do tempo de resolução, os alunos deveriam responder a um formulário com perguntas abertas e na escala *Likert* de 1 a 5 (sendo 1 - Discordo Totalmente e 5 - Concordo Totalmente) criado no Google Forms sobre percepção acerca da experiência do dia. E por fim, enviar as suas resoluções pela plataforma educacional Aprender3 para correção.

Após a conclusão dos experimentos, os exercícios enviados pelos estudantes na plataforma Aprender3 foram analisados e corrigidos manualmente. As correções realizadas foram organizadas em planilhas separadas, uma para cada dia de experimento, e armazenadas no Google Drive. Com o intuito de garantir uma correção consistente e padronizada, foram adotados alguns critérios para mensurar o desempenho dos alunos:

- **Execução do código:** O código roda corretamente sem erros?
- **Casos de teste:** O código acerta todos os casos de teste?
- **Indentação:** O código apresenta problemas na indentação?
- **Lógica do programa:** A lógica implementada está correta e condiz com o esperado?

¹Plataforma para resolução de problemas. Disponível em: <https://beecrowd.com/pt/>. Acesso em: 25 jan. 2025.

Além desses critérios de avaliação, foram identificados e classificados os tipos mais comuns de erros encontrados durante a correção dos exercícios. Esses erros foram agrupados nas categorias presentes na Tabela 1.

Tabela 5.1: Descrição dos tipos de erros.

Id	Tipo de Erro	Descrição
E01	Erro na coleta de entrada	Problemas relacionados à obtenção dos dados de entrada.
E02	Erros de lógica	Falhas na implementação da lógica necessária para resolução do problema proposto.
E03	Erros de formatação de saída	Falhas nas convenções de formatação esperadas.
E04	Erro de referência	Uso incorreto de variáveis ou referências.
E05	Erro de tipo	Incompatibilidade de tipos de dados durante a execução do programa
E06	Erro de sintaxe	Violação da sintaxe da linguagem Python.
E07	Erro de indexação	Problemas com o uso incorreto de índices em estruturas de dados.
E08	Erro de indentação	Problemas relacionados ao uso de espaços ou tabulações.

Capítulo 6

Perfil da Turma

Neste capítulo é apresentado o perfil geral da turma, mostrando a distribuição de gênero, bem como a experiência prévia dos estudantes com a programação e ferramentas de inteligência artificial, especificamente os *chatbots* como ChatGPT, Gemini, e etc.

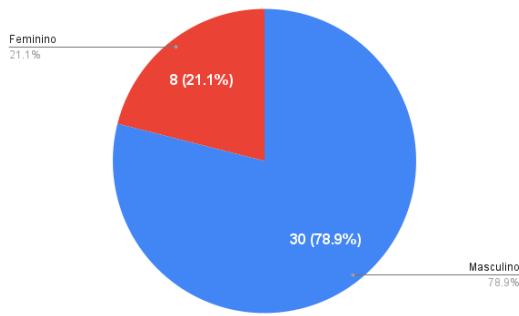
Para compreender o perfil dos estudantes matriculados na disciplina de Algoritmos e Programação de Computadores no semestre de 2024/2, do curso de Licenciatura em Computação, foi aplicado um formulário do Google Forms para os estudantes. O objetivo deste formulário foi mapear as características da turma para realizar a divisão dos estudantes em dois grupos (experimental e controle), garantindo que a composição dos grupos não comprometesse os resultados do estudo, visto que a pesquisa avalia o desempenho dos dois grupos em atividades realizadas em sala de aula, sendo que o grupo experimental terá acesso ao ChatGPT, enquanto o grupo controle não o utilizará. Ao total 38 dos 41 alunos responderam ao formulário.

Dessa forma, na Figura 6.1, os dados coletados revelaram que apenas 3 alunos estão cursando APC pela segunda vez ou mais, enquanto que a maioria (35 estudantes) está matriculada na disciplina pela primeira vez. Além disso, 25 alunos já tiveram algum contato prévio com programação antes de iniciar o curso, indicando um nível básico de familiaridade com os conceitos iniciais da disciplina. Por último, 29 estudantes declararam já ter utilizado ferramentas de inteligência artificial para auxiliar na programação, sugerindo que grande parte da turma tem alguma experiência com tecnologias de IA.

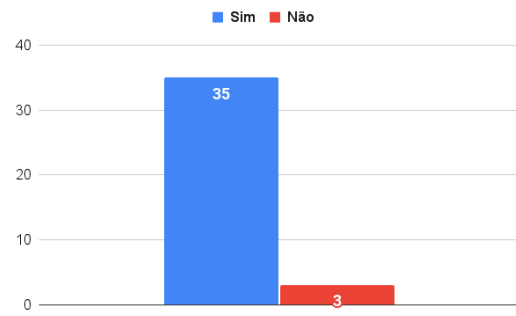
Com esses dados, é possível compreender melhor o perfil da turma na tentativa de mitigar possíveis vieses nos resultados do estudo.

6.1 Divisão dos Grupos

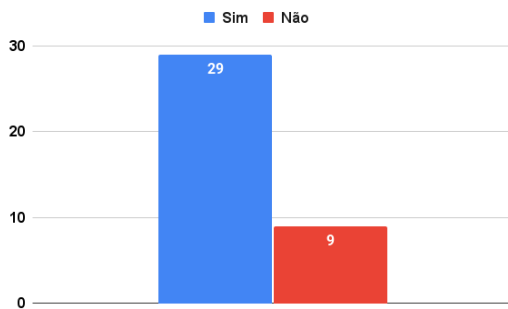
A fim de equiparar o perfil dos dois grupos, a separação da turma não será totalmente aleatória. Para isso, foram considerados, na divisão, os dados do perfil de cada um dos



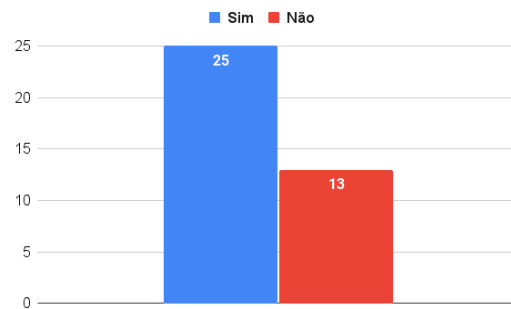
(a) Distribuição por sexo na turma.



(b) Quantidade de alunos cursando APC pela primeira vez.



(c) Quantidade de alunos que utilizaram assistentes virtuais anteriormente.



(d) Quantidade de alunos com experiência prévia em programação.

Figura 6.1: Dados referentes ao perfil dos alunos coletados durante a pesquisa.

38 alunos que responderam ao formulário mencionado anteriormente. Dessa forma, foram levados em conta os seguintes aspectos:

- Familiaridade do aluno com a programação;
- Habilidade do aluno em identificar e corrigir erros em códigos;
- Confiança do aluno em resolver erros e problemas de código sem assistência;
- Familiaridade do aluno com a ferramenta do ChatGPT;
- Autopercepção do aluno sobre a sua habilidade de resolver problemas de lógica.

Com base nesses critérios, os alunos foram distribuídos de forma a equilibrar os níveis de habilidade e conhecimento entre os grupos, evitando que um grupo tivesse alguma vantagem sobre o outro. Por exemplo, a divisão impediu que um grupo fosse formado apenas por alunos sem experiência alguma em programação, enquanto o outro contasse exclusivamente com alunos que já possuíam conhecimento prévio. Da mesma forma, também foi evitado que um grupo fosse formado apenas por alunos confortáveis em realizar tarefas sem assistência, enquanto o outro reunisse apenas aqueles que preferiam algum tipo de suporte.

Por último, três alunos que participaram do experimento não responderam ao formulário, e por isso, foram alocados aleatoriamente no primeiro dia em que estes participaram do experimento.

Capítulo 7

ChatGPT x Tutoria Tradicional

Este capítulo apresenta a análise dos resultados das atividades realizadas ao longo de três dias de experimento, considerando os dois grupos: experimental, representado pelos alunos que tiveram a assistência do ChatGPT como tutor em sala de aula, e o grupo controle, assistido pela equipe de monitoria presente em sala. A análise baseia-se nas respostas submetidas pelos estudantes na plataforma Aprender3, após as correções manuais. Nesta seção será avaliado o desempenho de cada grupo e comparativos dos resultados obtidos para verificar se o uso do ChatGPT influencia o desempenho acadêmico dos estudantes.

7.1 Participação e Desempenho dos Alunos nos Experimentos

A relação da quantidade de alunos em cada grupo e os conteúdos aos quais pertenciam os exercícios ministrados em aula durante a semana estão na Tabela 7.1. Ao longo dos experimentos, 41 alunos da turma participaram, sendo 20 no grupo controle e 21 no grupo experimental. Destaca-se, ainda, que 19 alunos estiveram presentes em todos os experimentos, dos quais 6 pertenciam ao grupo controle e 13 ao grupo experimental.

Tabela 7.1: Distribuição semanal dos conteúdos ministrados e quantidade de alunos presentes nos grupos controle e experimental

Semana	Data do Experimento	Conteúdos dos exercícios	Quantidade de alunos presentes em sala no grupo controle	Quantidade de alunos presentes em sala no grupo experimental	Total de alunos presentes no experimento
1	04/12/2024	Condicionais e Iteração	18	18	36
2	11/12/2024	Condicionais e Iteração	11	19	30
3	18/12/2024	Strings, Condicionais, Iteração e Listas	9	17	26

Nota-se que, ao longo das semanas, a quantidade de alunos presentes em sala foi diminuindo. No grupo controle, a taxa de ausência foi superior ao grupo experimental, chegando a menos da metade dos alunos pertencentes ao grupo no último dia de experimentos. Porém, é importante ressaltar que o período do ano em que o experimento foi realizado, mais especificamente a última semana, antecedeu o recesso de festas de fim de ano, o que pode ter contribuído para a baixa presença nesta semana.

Assim, em relação ao desempenho durante as três semanas em que o experimento foi realizado, nesta seção será apresentado um comparativo entre os grupos controle e experimental quanto ao término e ao acerto dos exercícios aplicados em cada dia; e por último, uma análise geral na subseção 7.5. Ressalta-se, também, que os erros comentados ao decorrer da análise estão identificados na Tabela 5.1.

7.2 Análise do Dia 1

No primeiro dia de experimento, os alunos foram apresentados à metodologia que seria aplicada e separados formalmente, pela primeira vez, nas duas salas. Na Tabela 7.2, são apresentadas as relações entre os grupos controle e experimental com base em dados nos resultados dos exercícios aplicados em sala de aula. A partir da análise de desempenho dos alunos neste experimento, considerando tanto o término quanto a correção dos exercícios, observa-se que o grupo experimental entregou e acertou mais exercícios do que o grupo controle. Além disso, em relação aos alunos que entregaram todos os exercícios propostos, o único aluno do grupo controle que completou todos acertou todas as questões, enquanto no grupo experimental, metade dos alunos que entregaram todos os exercícios acertaram todas as questões.

Tabela 7.2: Comparativo entre os grupos controle e experimental quanto ao término e acerto dos exercícios aplicados na Dia 1

	Grupo Controle	Grupo Experimental	Geral
Alunos que terminaram todos os exercícios	5.56%	33.33%	19.44%
Alunos que terminaram e acertaram todos os exercícios	100.00%	50.00%	57.14%
Relação entre o número de resoluções entregues e o total de exercícios que deveriam ser entregues	49.12%	66.67%	57.66%
Proporção entre o número de resoluções corretas e o total de exercícios submetidos	57.14%	72.22%	65.63%

Em relação aos erros, na Figura 7.1, foram contabilizados os erros cometidos e seus respectivos tipos. Observa-se que a maioria dos erros cometidos por ambos os grupos foi de lógica, na coleta de entrada, de referência e de indentação. Além disso, ambos os grupos classificaram a atividade como de nível médio. Quanto à utilização de tutoria (humana

Tabela 7.3: Comparativo entre os grupos controle e experimental quanto ao término e acerto dos exercícios aplicados na Dia 2

	Grupo Controle	Grupo Experimental	Geral
Alunos que terminaram todos os exercícios	36.36%	73.33%	23.08%
Alunos que terminaram e acertaram todos os exercícios	25.00%	45%	57.69%
Relação entre o número de resoluções entregues e o total de exercícios que deveriam ser entregues	72.73%	88.89%	82.05%
Proporção entre o número de resoluções corretas e o total de exercícios submetidos	66.67%	72.50%	70.31%

ou ChatGPT), 83.33% dos alunos do grupo controle consultaram os monitores durante a aula, enquanto 77.78% dos alunos do grupo experimental consultaram o ChatGPT para esclarecer dúvidas.

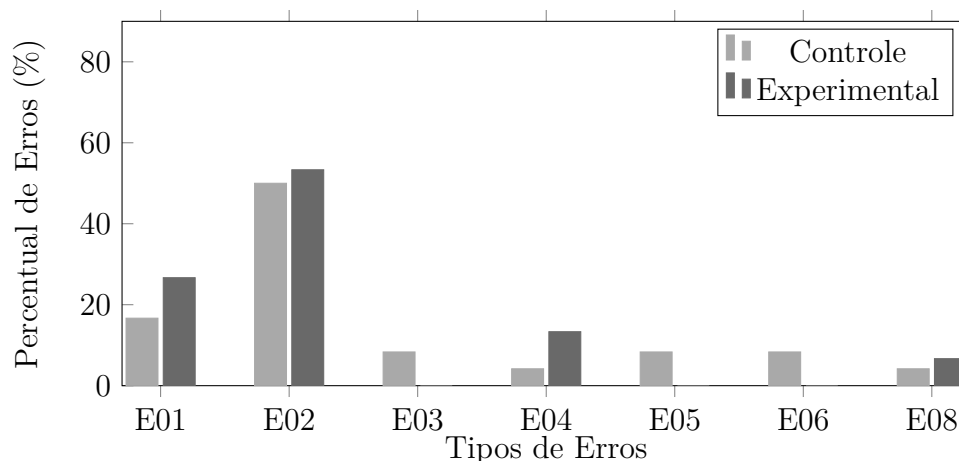


Figura 7.1: Percentual de Erros por Tipo de Erro - Comparação entre os grupos Controle e Experimental no Dia 1

7.3 Análise do Dia 2

No segundo dia de experimento, foram realizados exercícios sobre os mesmos conteúdos da atividade da semana anterior, pois no dia seguinte os alunos teriam a primeira prova da disciplina. Assim, como forma de revisão, o conteúdo dos exercícios foi o mesmo da prova 1.

Na Tabela 7.3, são apresentadas relações semelhantes às da Tabela 7.2, na Seção 7.2. Novamente, o grupo experimental se destacou em todas as métricas comparadas ao grupo controle. Ainda, pode-se observar, que de forma geral, houve uma melhora no desempenho dos dois grupos em comparação com a semana anterior.

Os alunos classificaram esse dia de atividades como de nível médio. Em relação aos erros, na Figura 7.2, foram contabilizados os erros cometidos e seus respectivos tipos. Observa-se que a maioria dos erros cometidos por ambos os grupos foi de lógica e formatação de saída. No entanto, destaca-se que o grupo controle cometeu, além desses, erros como: erro de referência, erro de sintaxe, erro de indexação e erro de indentação.

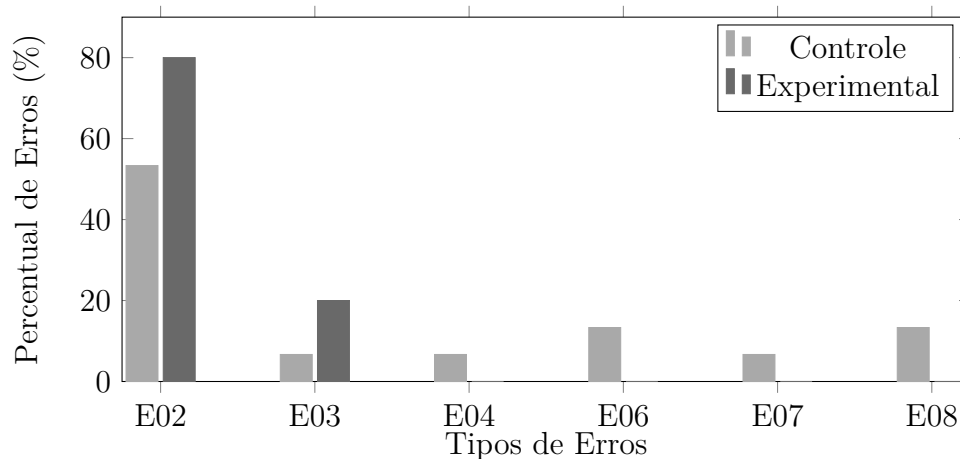


Figura 7.2: Percentual de Erros por Tipo de Erro - Comparação entre os grupos Controle e Experimental no Dia 2

Quanto à utilização de tutoria (humana ou ChatGPT), 90.91% dos alunos do grupo controle consultaram os monitores durante a aula, enquanto 89.47% dos alunos do grupo experimental consultaram o ChatGPT para esclarecer dúvidas.

Esse foi o dia com melhor desempenho geral da turma: mais alunos entregaram os exercícios e houve um maior número de acertos. Foi também o dia de experimento com melhor desempenho do grupo controle. Nota-se que, por ser uma aula na véspera da prova, os alunos possivelmente estavam mais focados em terminar os exercícios com o intuito de revisar e fixar os conteúdos. Esse fator pode ter sido determinante para a alta quantidade de exercícios entregues e maior busca por monitoria.

7.4 Análise do Dia 3

Nota-se que este último experimento aconteceu após a primeira prova. Além disso, a aula foi realizada na semana que antecedeu o recesso de festas de final de ano, o que pode ter contribuído para a baixa presença neste dia.

Assim, como a prova já havia ocorrido, foram apresentados mais conteúdos ministrados até o momento da aula nas atividades propostas. Esse fator pode ter contribuído para que a maioria da turma classificasse os exercícios como de nível difícil. Todavia, mesmo com essa classificação, a maior parte da turma entregou todos os exercícios. No entanto,

Tabela 7.4: Comparativo entre os grupos controle e experimental quanto ao término e acerto dos exercícios aplicados na Dia 3

	Grupo Controle	Grupo Experimental	Geral
Alunos que terminaram todos os exercícios	20.00%	60.00%	44.00%
Alunos que terminaram e acertaram todos os exercícios	0.00%	13.33%	8.00%
Relação entre o número de resoluções entregues e o total de exercícios que deveriam ser entregues	50.00%	80.00%	68.00%
Proporção entre o número de resoluções corretas e o total de exercícios submetidos	53.33%	61.11%	58.82%

menos de 10.00% conseguiu acertar completamente os exercícios entregues, como pode-se observar na Tabela 7.4, ressalta-se que nenhum aluno do grupo controle acertou todos os exercícios.

Em relação aos erros cometidos, na Figura 7.3 observa-se que os tipos mais recorrentes entre ambos os grupos foram os de lógica, formatação de saída e coleta de entrada. No entanto, também foram observados outros erros, como de referência, sintaxe, indexação e indentação.

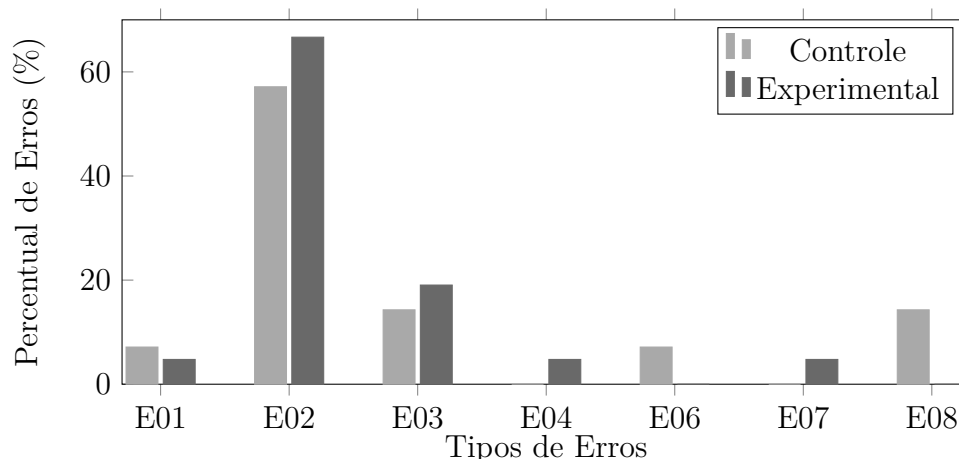


Figura 7.3: Percentual de Erros por Tipo de Erro - Comparação entre os grupos Controle e Experimental no Dia 3

Por fim, quanto a utilização de tutoria (humana ou ChatGPT), 66.67% dos alunos do grupo controle consultaram os monitores durante a aula, enquanto 100% dos alunos do grupo experimental consultaram o ChatGPT para dúvidas, ou seja, todos os alunos do grupo experimental utilizaram o ChatGPT neste dia.

Nota-se que o desempenho do grupo controle foi novamente mais baixo que o do grupo experimental. Mantendo o padrão encontrado entre os grupos nos dias anteriores de experimento.

7.5 Análise Geral

Sobre um desempenho comparativo entre o grupo controle e experimental, foi analisada a relação entre o número de resoluções entregues e total de resoluções esperadas, dados presentes na Tabela 7.5. Observa-se que o grupo experimental entregou significativamente mais resoluções em comparação com o grupo controle, superando consistentemente as médias gerais da turma em todos os dias analisados.

Tabela 7.5: Relação entre o número de resoluções entregues e o total de questões que deveriam ser resolvidas em cada dia de experimento

	Dia 1 (%)	Dia 2 (%)	Dia 3 (%)	Média Geral (%)	Desvio Padrão
Grupo Controle	49.12	72.73	50.00	57.28	13.38
Grupo Experimental	66.67	88.89	80.00	78.52	11.18
Geral	57.66	82.05	68.00	69.24	12.24
Desvio Padrão	8.78	8.11	15.10	10.65	-

Agora, em relação ao número de resoluções corretas e ao total de resoluções enviadas, observa-se uma diferença bem menor entre os valores obtidos dos dois grupos, que apresentam resultados próximos à média da turma na maioria dos experimentos, conforme é apresentado na Tabela 7.6. Entretanto, é possível observar uma superioridade dos resultados obtidos pelo grupo experimental em relação ao grupo controle. Essa superioridade foi mais evidente no Dia 1, com uma diferença bastante alta, em que o desvio padrão foi de 7.56 entre os grupos e o desempenho geral da turma. Porém, nos dias restantes (Dias 2 e 3), essa diferença foi menor, com os desvios padrões variando entre 2.94 a 3.99, indicando um maior equilíbrio nos desempenhos.

Tabela 7.6: Proporção entre o número de resoluções corretas em relação as questões enviadas em cada dia do experimento

	Dia 1 (%)	Dia 2 (%)	Dia 3 (%)	Média Geral (%)	Desvio Padrão
Grupo Controle	57.14	66.67	53.33	59.05	6.87
Grupo Experimental	72.22	72.50	61.11	68.61	6.50
Geral	65.63	70.31	58.82	64.92	5.78
Desvio Padrão	7.56	2.94	3.99	4.82	-

Logo, pode-se notar que os alunos que utilizaram o ChatGPT entregaram mais exercícios do que os alunos que não utilizaram. No entanto, isso não implica necessariamente que eles tenham acertado mais exercícios. Dessa forma, pode-se sugerir que o volume de resoluções entregues não é diretamente proporcional à precisão ou qualidade das respostas fornecidas. Ou seja, embora o uso do ChatGPT tenha permitido que os alunos entregassem mais resoluções, o impacto sobre a qualidade das respostas não se mostrou significativamente diferente em relação aos alunos que utilizaram a monitoria de alunos.

Outro ponto analisado foi quanto a dificuldade dos exercícios. De acordo com dados presentes nas Figuras 7.4, 7.5 e 7.6, os dois grupos classificaram a atividade de forma semelhante, apresentando distribuições parecidas em todos os dias em que esse quesito foi avaliado. Por último, a afirmação “Tive mais dúvidas quanto ao enunciado dos exercícios do que em relação ao código a ser gerado.”, busca compreender a percepção dos alunos sobre a origem de suas dificuldades.

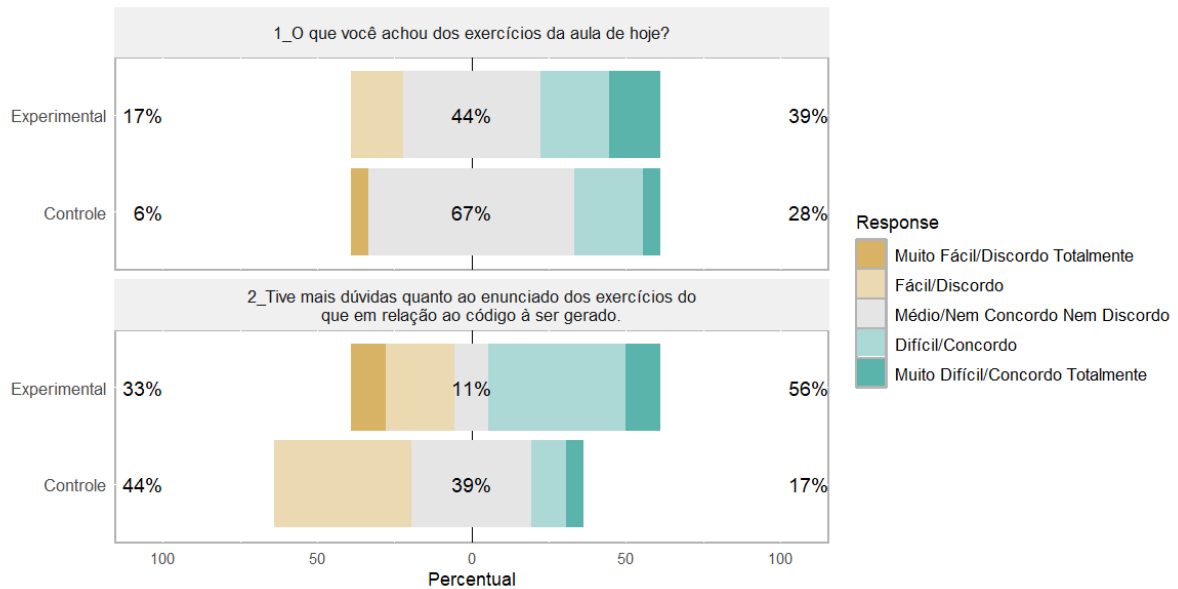


Figura 7.4: Comparação de dificuldade e dúvidas - Dia 1.

Nesse quesito, houve mais divergência entre os grupos: enquanto o grupo controle, no primeiro dia, apontou que possuía mais dificuldade com a codificação, o grupo experimental apresentou maior dificuldade na compreensão do enunciado. No segundo dia, o grupo controle afirmou que a maior dificuldade estava na compreensão do enunciado, enquanto o grupo experimental se dividiu entre neutralidade e dificuldades na compreensão do enunciado. No último dia, o grupo controle apresentou maior dificuldade com a codificação e uma parcela com neutralidade, enquanto, no grupo experimental, a maioria encontrou dificuldades na codificação. No entanto, alguns alunos indicaram dificuldades na compreensão do enunciado, além de certa neutralidade. Ou seja, dependendo da atividade, dificuldades diferentes foram enfrentadas pelos grupos, de acordo com a ferramenta à disposição.

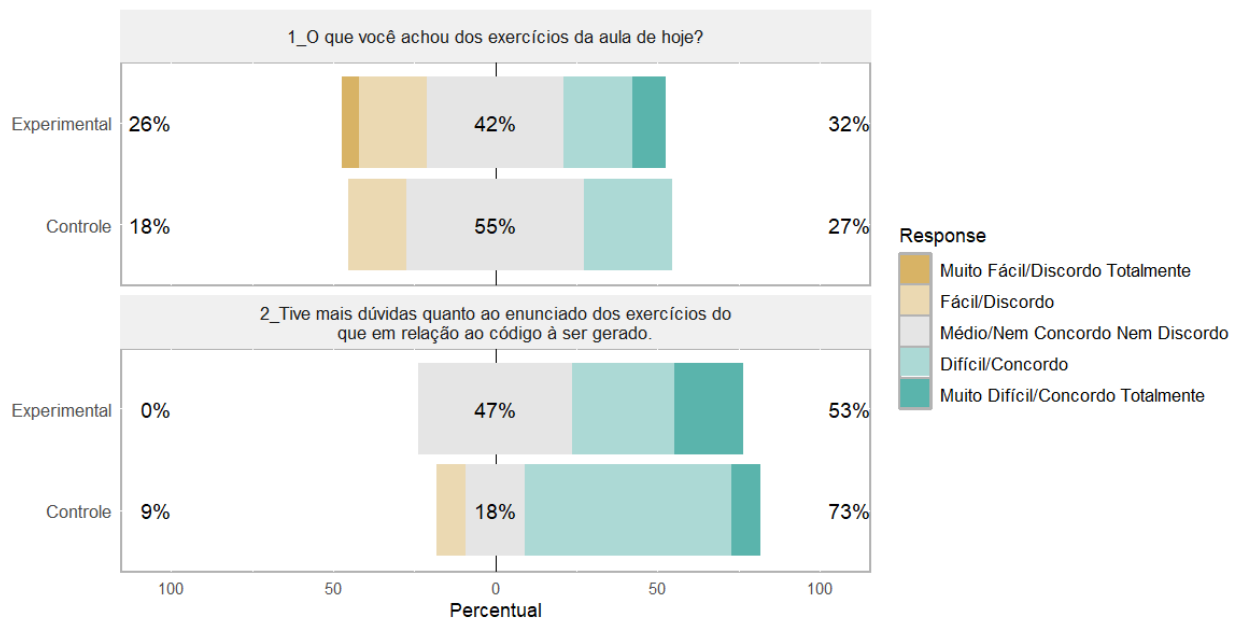


Figura 7.5: Comparação de dificuldade e dúvidas - Dia 2.

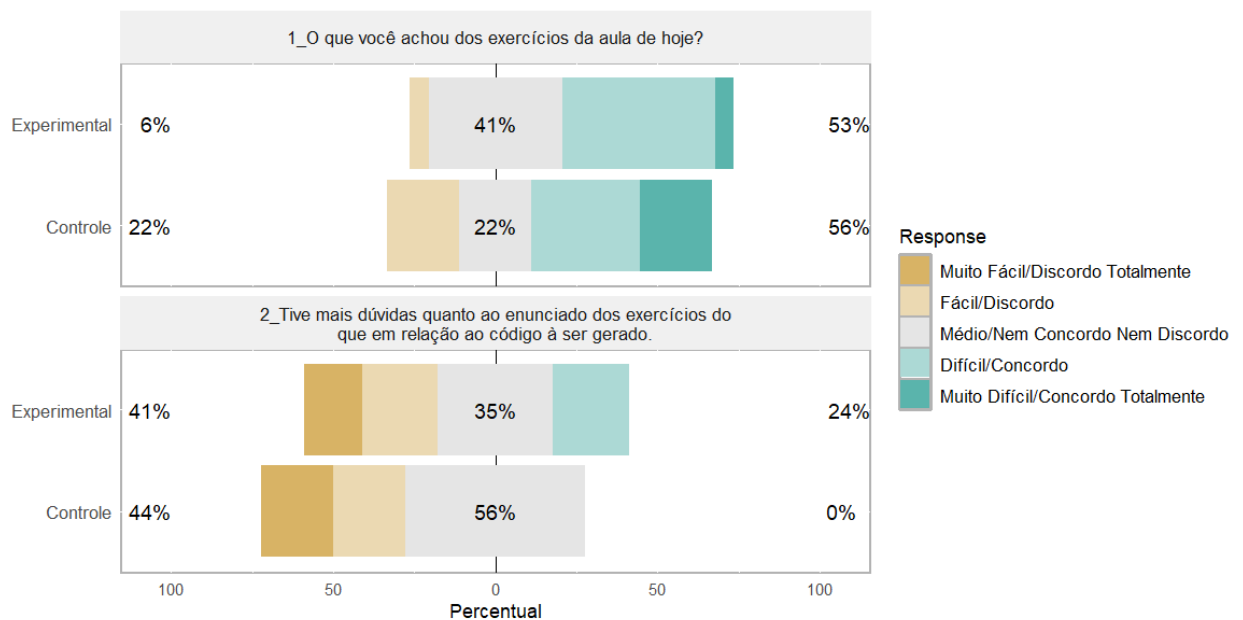


Figura 7.6: Comparação de dificuldade e dúvidas - Dia 3.

Capítulo 8

Percepção dos estudantes sobre o uso do ChatGPT

Neste capítulo, são apresentadas as respostas dos estudantes do grupo experimental ao formulário de pós-atividade para compreender as percepções dos participantes sobre o uso do ChatGPT nas atividades em laboratório. A Figura 8.1 e a Tabela 8.1 apresentam as perguntas de cunho geral, às quais tanto os alunos do grupo experimental que usaram o ChatGPT quanto os alunos do grupo experimental que optaram por não utilizá-lo responderam. Já a Figura 8.2 e a Tabela 8.2 apresentam a categoria de perguntas relacionadas aos sentimentos sobre o uso da ferramenta, que foram respondidas apenas pelos alunos do grupo experimental que utilizaram o ChatGPT para realizar as atividades em sala.

8.1 Percepções do Dia 1

Para o primeiro dia de experimento, obteve-se ao total 18 respostas dos participantes no formulário de pós-atividade para questões de sentimento geral e ao total 14 respostas para as questões de sentimento sobre o uso da ferramenta durante a atividade (veja as Tabelas 8.1 e 8.2). Ao analisar as respostas dos participantes ao formulário, observa-se na Figura 8.1 que a maioria dos estudantes, 93%, concorda que o ChatGPT é uma ferramenta amigável (Q1). Este resultado pode indicar que a interface e o funcionamento da ferramenta foram bem aceitos pelos alunos. Na Figura 8.2, a maioria dos participantes, ou seja, 79%, concordou que utilizaria o ChatGPT fora da sala de aula para auxiliar na programação (Q14), sugerindo que os participantes consideraram a ferramenta útil.

Além disso, na Figura 8.2, 71% dos alunos concordaram que o ChatGPT foi capaz de esclarecer dúvidas de forma eficiente (Q10), o que pode sugerir seu potencial como suporte ao aprendizado. Por outro lado, na Figura 8.1, apenas 57% dos participantes concordaram que o ChatGPT consegue identificar erros e *bugs* de forma rápida e eficiente (Q3). Na

Tabela 8.1: Resultados referentes às questões de sentimento geral sobre o ChatGPT no formulário pós-atividade.

Sentimento Geral		Dia 1 (n = 18)				Dia 2 (n = 19)				Dia 3 (n = 17)			
		M	DP	Me	Mo	M	DP	Me	Mo	M	DP	Me	Mo
Q1	ChatGPT é uma ferramenta amigável.	4.22	1.00	4	5	4.05	0.97	4	5	3.88	0.78	4	4
Q2	ChatGPT pode substituir completamente a equipe de monitoria em qualquer situação relacionada à programação na sala de aula.	-	-	-	-	1.73	0.80	2	1	1.64	0.78	1	1
Q3	ChatGPT pode identificar erros e bugs de forma rápida e eficiente.	3.38	1.03	3.5	4	3.52	1.02	4	4	3.82	0.63	4	4
Q4	ChatGPT é muito bom em programação.	-	-	-	-	3.26	1.09	3	4	3.41	0.71	4	4
Q5	Eu confio nas respostas do ChatGPT para questões de programação.	2.66	1.08	3	3	2.84	0.83	3	3	3.05	0.65	3	3
Q6	O uso do ChatGPT deveria ser permitido nas atividades de APC.	-	-	-	-	3.57	1.34	3	3	3.58	1.17	3	3
Q7	Tive mais dúvidas quanto ao enunciado dos exercícios do que em relação ao código à ser gerado.	3.22	1.26	4	4	3.73	0.80	4	3	2.64	1.05	3	3
Q8	Prefiro tirar dúvidas com o ChatGPT do que com os monitores.	-	-	-	-	2.31	1.00	2	2	2.17	0.80	2	3

Nota: n representa a quantidade de estudantes que responderam a estas questões.

Legenda: M = Média, DP = Desvio Padrão, Me = Mediana, Mo = Moda.

Figura 8.2, 64%, acreditou que o suporte do ChatGPT foi suficiente para ajudá-los a encontrar erros e desenvolver códigos (Q11).

Quanto à questão (Q9) sobre o ChatGPT ser crucial para a realização das atividades, as respostas dos alunos estão bem divididas: 43% mantiveram uma opinião neutra e 43% concordaram que ela foi indispensável (Figura 8.2). Esses resultados indicam que, para uma parcela significativa dos alunos, o uso do *chatbot* não foi essencial, mas também não foi irrelevante. Por outro lado, para a outra parcela significativa, o uso da ferramenta foi indispensável, indicando que parte dos alunos viu a ferramenta como um recurso fundamental para concluir as atividades.

Na Figura 8.2, apenas 21% dos estudantes acreditam que por ter utilizado o ChatGPT aprenderam mais em comparação com o grupo controle (Q13), enquanto uma proporção significativa discorda (43%) ou mantém uma opinião neutra (36%). Cerca de 50% dos estudantes apresentou uma opinião neutra sobre a confiabilidade das respostas do ChatGPT para questões de programação (Q5), seguida por 36% que discordaram e apenas 14% que concordaram (Figura 8.1). Esses resultados sugerem uma preocupação dos estudantes em relação à confiabilidade da ferramenta.

Dezoito alunos do grupo experimental estavam presentes no laboratório. Desses, 14 utilizaram o ChatGPT e 4 optaram por não utilizá-lo, ou seja, 77.78% fizeram uso da ferramenta. Os participantes Aluno 11, Aluno 6, Aluno 33 e Aluno 28 optaram por não utilizar o ChatGPT no dia 1.

Dessa maneira, considerou-se interessante acrescentar novas questões ao formulário para o segundo dia de experimento. Uma delas seria voltada a entender o motivo pelo

Tabela 8.2: Resultados referentes às questões de sentimento sobre o uso do ChatGPT no formulário pós-atividade.

Sentimento Uso na Atividade		Dia 1 (n = 14)				Dia 2 (n = 17)				Dia 3 (n = 17)			
		M	DP	Me	Mo	M	DP	Me	Mo	M	DP	Me	Mo
Q9	Acredito que o uso do ChatGPT foi crucial para eu realizar a tarefa.	3.35	0.84	3	3	2.94	0.96	3	3	3.76	0.90	4	4
Q10	O ChatGPT foi eficaz e eficiente ao tirar minhas dúvidas.	3.85	1.02	4	4	3.70	1.04	4	4	3.88	0.69	4	4
Q11	Acredito que o suporte fornecido pelo ChatGPT foi suficiente para eu encontrar os erros, e desenvolver os códigos necessários.	3.64	1.08	4	4	3.35	1.27	4	4	3.88	0.60	4	4
Q12	O ChatGPT parece responder mais rápido às dúvidas em comparação a equipe de monitoria na sala de aula.	-	-	-	-	3.11	1.21	3	3	3.11	0.78	3	3
Q13	Acredito que, por ter utilizado o ChatGPT, aprendi mais do que o outro grupo.	3.00	1.17	3	2	2.76	1.20	3	3	2.64	1.27	3	1
Q14	Fora dos experimentos, continuarei a utilizar o ChatGPT para me auxiliar na programação.	4.21	0.80	4	5	3.88	1.05	4	5	3.82	1.23	4	5

Nota: n representa a quantidade de estudantes que responderam a estas questões.

Legenda: M = Média, DP = Desvio Padrão, Me = Mediana, Mo = Moda.

qual alguns participantes optaram por não utilizar o ChatGPT.

8.2 Percepções do Dia 2

Para o segundo dia de experimento, o formulário de pós-atividade obteve a participação de 19 estudantes para questões de sentimento geral e ao total 17 respostas para as questões de sentimento sobre o uso da ferramenta durante a atividade (veja as Tabelas 8.1 e 8.2). Ao analisar as respostas dos participantes ao formulário na Figura 8.1, observar-se que a maioria dos estudantes, 68%, concordaram que o ChatGPT é uma ferramenta amigável (Q1). Com isso, nota-se que houve uma diminuição significativa do primeiro dia que foi de 93% para o segundo dia. Houve um aumento significativo na porcentagem de alunos com opinião neutra em relação a essa sentença, passando de 0% no primeiro dia para 26% no segundo. Apesar disso, é possível concluir que a ferramenta manteve uma boa aceitação entre os alunos.

Quanto à questão sobre o ChatGPT conseguir identificar erros e *bugs* de forma rápida e eficiente (Q3), houve um pequeno aumento na porcentagem de alunos com opinião positiva em relação à afirmação, passando de 57% no primeiro dia para 63% no segundo dia.

Uma nova questão acrescentada buscou verificar se os participantes acreditam que o uso do ChatGPT deveria ser permitido nas atividades de APC (Q6). Verifica-se que a porcentagem de respostas que concordam (47%) está próxima das respostas com opinião neutra (37%). A proximidade entre as porcentagens de respostas que concordam e as respostas neutras pode significar que os estudantes têm opiniões divididas ou estão incertos

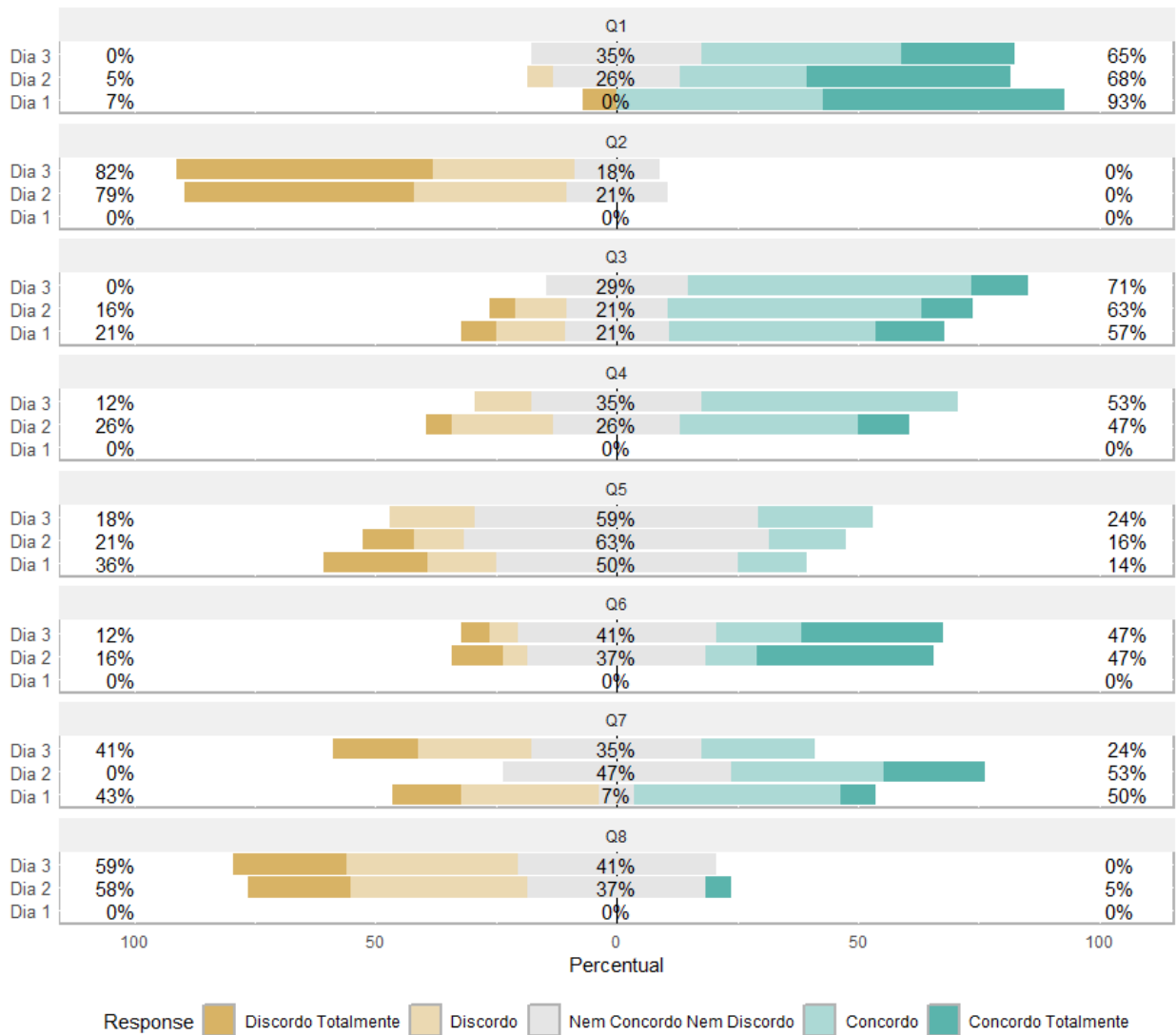


Figura 8.1: Sentimento geral sobre o ChatGPT.

quanto à inclusão do ChatGPT nas atividades de APC. A maioria, 47%, dos estudantes concordou que o ChatGPT é muito bom em programação (Q4). No entanto, observa-se que os alunos que discordaram e os que possuem uma opinião neutra estão igualmente divididos, ambos com 26%.

Já para a questão sobre confiabilidade nas respostas do ChatGPT para questões de programação (Q5) teve-se um aumento considerável de opiniões neutras saindo de 50% no primeiro dia para 63% no segundo dia. Houve também uma diminuição considerável na porcentagem de respostas que discordam saindo de 36% no primeiro dia para 21% no segundo dia.

A maioria dos participantes discorda (58%) da ideia de preferir tirar dúvidas com o ChatGPT ao invés de monitores (Q8). Em contrapartida, 37% tem uma opinião neu-

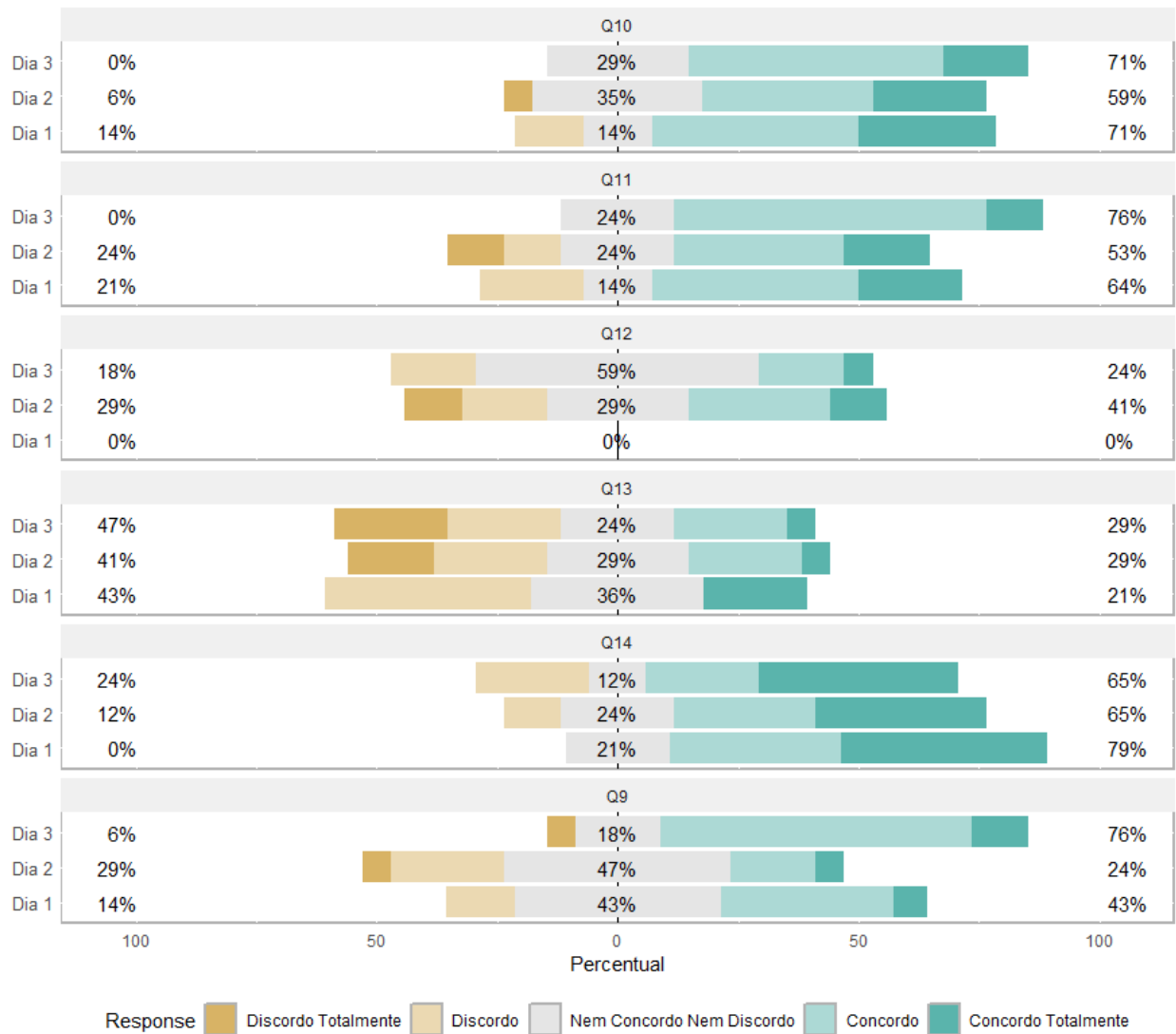


Figura 8.2: Sentimento sobre o uso do ChatGPT na atividade em sala de aula.

tra. Além disso, uma parcela ainda maior dos participantes é contrária à ideia de que o ChatGPT pode substituir completamente a equipe de monitoria em atividades relacionadas à programação em sala de aula (Q2), 79% discordaram, 21% com opiniões neutras e 0% concordaram. Dessa forma, pode-se sugerir que, embora o ChatGPT seja considerada uma ferramenta útil pelos alunos, eles ainda consideram importante interações humanas com monitores e tutores para sanar suas dúvidas.

Na Figura 8.2, observa-se que a maioria dos participantes, 65%, concordaram que continuarão utilizando o ChatGPT para o auxílio na programação (Q14). Dessa forma, houve uma diminuição na porcentagem de alunos que concordaram no primeiro dia passando de 79% para 65% no segundo dia.

A maioria dos alunos, 59%, concordou que o ChatGPT foi eficaz e eficiente ao tirar

dúvidas (Q10), uma diminuição significativa comparado ao primeiro dia de 71%. Além disso, 53% acreditaram que o suporte do ChatGPT foi suficiente para ajudá-los a encontrar erros e desenvolver códigos (Q11), uma redução, quando 64% dos alunos concordaram com essa mesma afirmação no primeiro dia.

Quanto à questão sobre o ChatGPT responder mais rapidamente às dúvidas em comparação com a equipe de monitoria em sala de aula (Q12), a maioria dos participantes, 41%, concordou. Por outro lado, as opiniões estão divididas entre aqueles que discordaram e os que tiveram uma opinião neutra, ambos com 29%.

Em relação à afirmação “por ter utilizado o ChatGPT, aprendi mais que o outro grupo”(Q13), a maioria dos participantes, 41%, discordou. Já as opiniões neutras e as que concordaram estão igualmente divididas, ambas com 29%. Por fim, quanto à questão sobre o ChatGPT ser crucial para a realização das atividades (Q9), a maioria apresentou uma opinião neutra (47%). As respostas dos alunos que concordaram e discordaram estão divididas: 24% concordaram que a ferramenta foi crucial, enquanto 29% discordaram de sua indispensabilidade.

Dezenove alunos do grupo experimental estavam presentes no laboratório. Desses, 17 optaram por utilizar o ChatGPT na realização dos exercícios e 2 optaram por não utilizá-lo. Assim, 89,47% utilizaram a ferramenta. Os alunos Aluno 6 e Aluno 35 optaram por não utilizar o ChatGPT no dia 2.

Com isso, foi possível perceber uma possível maior aceitação por parte dos alunos em relação ao uso do ChatGPT. Verificou-se que os alunos Aluno 11, Aluno 33 e Aluno 28 passaram a utilizar o ChatGPT. Entretanto, o Aluno 6 continuou a não utilizar a ferramenta, e um novo aluno, que havia utilizado o ChatGPT no primeiro dia, optou por não utilizá-lo no segundo dia.

O Aluno 6 respondeu que não utilizou o ChatGPT porque prefere realizar as tarefas sozinho(a), sem auxílios externos. Similar ao Aluno 6, o Aluno 35 respondeu que prefere realizar as tarefas sozinho(a), sem auxílios externos, e também que prefere tirar dúvidas com a equipe de monitoria.

8.3 Percepções do Dia 3

Por fim, o terceiro dia de experimento, obteve-se ao total 17 respostas dos participantes no formulário de pós-atividade para ambas as categorias de questões (veja as Tabelas 8.1 e 8.2). Na Figura 8.1, verifica-se que a maioria dos estudantes, 71%, concorda que o ChatGPT consegue identificar erros e *bugs* de forma rápida e eficiente (Q3). Houve um aumento na porcentagem de alunos que concordaram em relação a essa afirmação, passando de 63% no segundo dia para 71% no terceiro dia. Houve também uma diminuição

significativa na porcentagem de alunos que discordaram, passando de 16% no segundo dia para 0% no terceiro dia.

Quanto à questão sobre o ChatGPT ser uma ferramenta amigável (Q1), houve uma pequena queda na porcentagem de alunos com opinião positiva em relação à afirmação, passando de 68% no segundo dia para 65% no terceiro dia, além de uma variação considerável de alunos com opinião neutra, que passou de 26% para 35%. A maioria dos estudantes, 53%, concordou que o ChatGPT é muito bom em programação (Q4) que anteriormente era de 47%. Observa-se também um aumento na porcentagem de participantes com opinião neutra passando de 26% para 35%.

Em relação à opinião dos participantes sobre a permissão do uso do ChatGPT nas atividades de APC (Q6), observa-se que as respostas permaneceram similares aos resultados encontrados no segundo dia. A maioria dos participantes concordou (47%), mantendo a mesma porcentagem registrada anteriormente. Já a porcentagem de estudantes com opinião neutra aumentou de 37% para 41%. Essa proximidade entre as porcentagens de respostas que concordam e as respostas neutras continua indicando que os participantes ainda têm opiniões divididas ou estão incertos quanto à inclusão do ChatGPT nas atividades de APC.

Com relação à questão sobre confiabilidade nas respostas do ChatGPT para questões de programação (Q5) teve-se uma pequena queda de opiniões neutras saindo de 63% no segundo dia para 59% no terceiro dia. Houve também um pequeno aumento na porcentagem de respostas que concorda saindo de 16% no segundo dia para 24% no terceiro dia.

A maioria dos participantes discorda (59%) da ideia de preferir tirar dúvidas com o ChatGPT ao invés da equipe de monitores (Q8), mantendo uma porcentagem similar à do segundo dia, que foi de 58%. Em contrapartida, 41% têm uma opinião neutra, valor que era de 37% no segundo dia. Além disso, uma parcela ainda maior dos participantes é contrária à ideia de que o ChatGPT pode substituir completamente a equipe de monitoria em atividades relacionadas à programação em sala de aula (Q2): 82% discordaram, comparado a 79% anteriormente; 18% mantêm opiniões neutras, contra 21% anteriormente; e 0% concordaram, permanecendo como antes. Esses dados continuam indicando a ideia de que, embora o ChatGPT seja considerado uma ferramenta útil, os alunos ainda valorizam a interação humana com monitores para esclarecer suas dúvidas.

Na Figura 8.2, observa-se que 76% consideraram o suporte fornecido pelo ChatGPT eficiente para ajudá-los a encontrar erros e desenvolver códigos (Q11), representando um aumento significativo em relação ao segundo dia, quando esse índice foi de 53%. Além disso, houve uma redução na porcentagem de alunos que discordaram dessa afirmação, passando de 24% no segundo dia para 0% no terceiro.

Em relação à questão sobre o uso do ChatGPT ser crucial para a realização da tarefa (Q9), observou-se um aumento significativo de 24% no segundo dia para 76% no terceiro dia entre os alunos que concordaram com essa afirmação. De forma paralela, houve uma diminuição significativa nas porcentagens de alunos com opiniões neutras ou discordantes, passando de 47% para 18% e de 29% para 6%, respectivamente.

No que diz respeito à eficácia e eficiência do ChatGPT em tirar dúvidas (Q10), 71% dos participantes concordaram, um aumento em comparação aos 59% registrados no segundo dia. Em relação à afirmação sobre o uso da ferramenta para auxílio na programação fora dos experimentos (Q14), o percentual de concordância manteve-se constante em 65%. No entanto, houve um aumento de 12% para 24% na porcentagem de participantes que discordaram dessa afirmação. Por outro lado, em relação à declaração “por ter utilizado o ChatGPT, aprendi mais do que o outro grupo” (Q13), a maioria dos participantes, 47%, discordou, o que representa um pequeno aumento em comparação aos 41% registrados no segundo dia.

Quanto à questão sobre o ChatGPT responder mais rapidamente às dúvidas em comparação com a equipe de monitoria em sala de aula (Q12), a maioria dos participantes, 59%, manteve uma opinião neutra, um aumento significativo em relação ao segundo dia (29%). As opiniões foram ligeiramente divididas entre aqueles que discordaram (18%) e os que concordaram (24%).

Dezessete alunos do grupo experimental estavam presentes no laboratório neste dia. Desses 17 alunos, todos optaram por utilizar o ChatGPT na realização dos exercícios. Assim, 100% dos estudantes fizeram uso da ferramenta. Dessa forma, percebe-se que, ao longo dos três dias de experimentos, a taxa de aceitação por parte dos alunos em relação ao uso do ChatGPT foi crescente. Ao analisar as respostas dos participantes, constatou-se que os alunos Aluno 11, Aluno 33 e Aluno 28 continuaram utilizando o ChatGPT desde o segundo dia. O Aluno 6, que havia optado por não utilizar a ferramenta no segundo dia, passou a utilizá-la no terceiro. Já o Aluno 35 faltou no último dia de experimento.

8.4 Percepções Gerais

Ao abordar os estudantes sobre questões como “Como você utilizou o ChatGPT?”, observa-se que, de maneira geral, o tema mais frequente, aparecendo em 41.93% das respostas, foi a busca por respostas a perguntas sobre conceitos. Essas perguntas variam desde funções básicas, como “como funciona a função *input?*”, até funções de ordenação, como o *sort* em Python. Empatados em segundo lugar, com 35.48% das respostas, estão os temas de exemplificação do enunciado e identificação de erros. Essa porcentagem representa as solicitações dos alunos para que o ChatGPT os ajudasse a identificar seus erros

Tabela 8.3: Perguntas abertas sobre percepção dos estudantes.

Perguntas	SIM	NÃO	TALVEZ
Você acredita que o ChatGPT pode ser útil em sala de aula?	93.75%	6.25%	-
Você acredita que o uso do ChatGPT em sala de aula pode causar algum impacto negativo?	66.66%	20.00%	13.33%

Para a pergunta “Você acredita que o uso do ChatGPT em sala de aula pode causar algum impacto negativo?”, foi possível observar que os alunos apresentaram diversas preocupações. A mais citada refere-se ao processo de aprendizagem; muitos alunos relataram a preocupação de que a ferramenta poderia prejudicar o desenvolvimento do raciocínio e da aprendizagem. Alguns comentários que merecem destaque são: “Sim, pois acredito que a dificuldade e a demora enfrentadas pelo aluno, sem o uso de ferramentas, são essenciais para o desenvolvimento da lógica.” e “Acredito que sim, o uso deliberado dele pode ocasionar na falsa sensação de aprendizagem.”.

A comodidade também foi uma preocupação recorrente entre os alunos, que temem que, pela facilidade e rapidez em obter respostas do ChatGPT, os estudantes deixem de pensar por conta própria e/ou copiem as respostas do ChatGPT como se fossem suas. Um comentário que se destaca entre os demais abordou um ponto diferente, mencionando que o ChatGPT pode fornecer informações excessivas para os alunos, o que pode causar confusão para iniciantes: “Sim, acredito que, no início da aprendizagem, o uso do ChatGPT pode ter um impacto negativo, considerando que ele pode apresentar informações mais avançadas do que o nível do aprendiz, causando confusão.”.

Outro ponto importante observado nas respostas foi a preocupação com a dependência excessiva do ChatGPT em atividades de programação. Um comentário relevante foi: “Talvez, ao não incentivar completamente o pensamento lógico durante a realização dos exercícios, o aluno não se esforce ao máximo antes de utilizar a ferramenta de inteligência artificial.”.

Capítulo 9

O uso do ChatGPT pode ter impacto na nota dos alunos na avaliação da disciplina?

Realizada por meio da plataforma Aprender3, a prova durou 150 minutos (2 horas e 30 minutos), e sua correção foi automática por meio do plug-in Code Runner ¹. Os conteúdos da disciplina foram abordados de maneira integrada nas três questões de prova. Durante a avaliação, os alunos poderiam somente consultar os slides disponíveis no Aprender3 da disciplina, não sendo permitidas dúvidas com os monitores sobre as questões da prova, e a avaliação deveria ser realizada individualmente.

Ao todo 35 alunos realizaram a prova, e a média geral da turma foi de 5.35. A porcentagem de aprovação geral foi de 52.94%, ou seja, cerca da metade da turma obteve nota superior à média. No que diz respeito aos grupos, a porcentagem de aprovação na prova foi similar, o grupo controle obteve 43.75% de aprovação, enquanto o experimental obteve 63.16%. A distribuição das menções foi conforme o Figura 9.1, em que MM, MS e SS são menções acima da média, e SR, II e MI são menções abaixo da média.

Lembrando que durante a prova nenhum dos alunos obteve nenhum auxílio externo. Para analisar a distribuição das notas entre os grupos, utilizou-se o Teste T para duas amostras independentes (*Independent-samples t-test*), no qual foi comparado o desempenho dos dois grupos (controle e experimental) com base nas notas e questões da prova. Sobre o *p-value* decorrente do *T-Test* aplicado: se o valor de *p-value* for maior ou igual a 0.05, significa que não há diferença estatisticamente significativa entre os grupos comparados; porém, se o *p-value* for menor que 0.05, significa que há uma diferença estatisticamente significativa entre os grupos.

¹“Code Runner” consiste em um plug-in de avaliação de respostas submetidas em relação a uma questão presente no ambiente do Aprender3. É capaz de avaliar códigos em relação a uma resolução esperada.

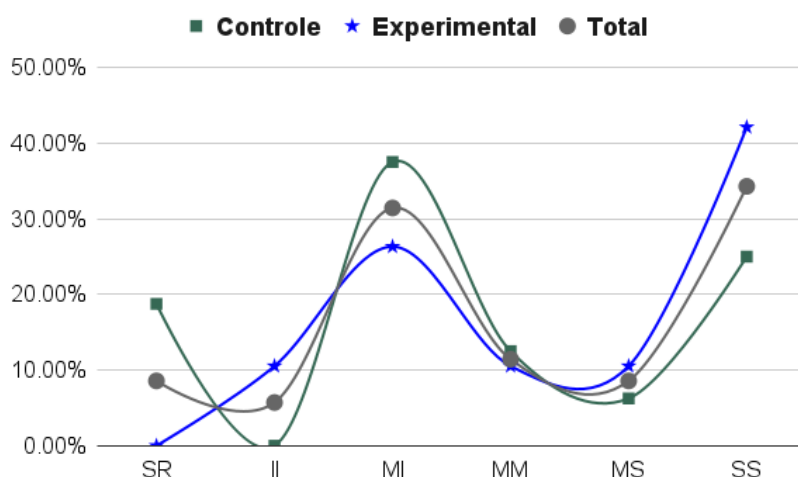


Figura 9.1: Distribuição das Menções da Prova

Tabela 9.1: Estatísticas Descritivas dos Resultados da Prova

	Média Grupo Controle	Desvio Padrão Grupo Controle	Média Grupo Experimental	Desvio Padrão Grupo Experimental	T-Test (p-value)
Questão 1 (3.00)	1.78	1.14	1.86	1.15	0.82
Questão 2 (4.00)	1.52	1.90	2.38	1.80	0.17
Questão 3 (3.00)	2.05	1.19	2.54	1.02	0.19
Prova (10.00)	5.35	3.51	6.80	3.13	0.20

Conforme apresentado na Tabela 9.1, o resultado obtido indicou que, em todos os âmbitos de análise estatística da prova, os dois grupos apresentaram resultados semelhantes. Portanto, não foi encontrado evidência estatística suficiente para afirmar que há diferenças significativas entre o desempenho do grupo controle e o grupo experimental na prova em questão, pois nenhum p -value ficou abaixo de 0.05.

Capítulo 10

Lições Aprendidas

Durante o período relatado, 41 alunos participaram da experiência, sendo 21 pertencentes ao grupo experimental, que utilizou o ChatGPT, e 20 ao grupo controle. Dessa forma, destacam-se as seguintes lições aprendidas:

- Inicialmente, nem todos os alunos do grupo experimental escolheram utilizar o ChatGPT, mesmo sendo permitidos e incentivados à utilizá-lo. Porém, ao longo dos três dias de experiência, mais alunos começaram a utilizar a ferramenta, mesmo anteriormente não tendo utilizado. Apesar do aumento progressivo da utilização do ChatGPT, em todos os dias avaliados os alunos relataram preferir a tutoria tradicional em comparação à ferramenta em questão.
- As opiniões dos alunos sobre a permissão do uso do ChatGPT em sala de aula foram divididas entre neutras e positivas. Divisão que reflete na preocupação dos participantes nas respostas abertas sobre o impacto negativo que a ferramenta poderia ter. Observou-se também que uma parcela significativa dos alunos possui uma posição neutra com relação a confiabilidade nas respostas do ChatGPT.
- A maioria dos alunos acredita que o ChatGPT pode ser uma ferramenta útil em sala de aula. Durante os dias de experiência, os alunos utilizaram a ferramenta para esclarecer dúvidas sobre conceitos, exemplificação do enunciado e identificação de erros.
- Os alunos que utilizaram do auxílio do ChatGPT foram capazes de entregar mais exercícios em menos tempo em relação aos alunos que utilizaram a monitoria ou não tiveram nenhum auxílio. No entanto, a proporção de acerto dos exercícios entre os dois grupos foi bastante similar. Assim, apesar do uso do ChatGPT permitir que os alunos entregassem mais exercícios em menos tempo, não necessariamente garantiam que as resoluções submetidas eram de maior qualidade ou precisão nas respostas.

- Durante a prova, os alunos não tinham acesso à tutoria tradicional ou ao ChatGPT. Os resultados obtidos pelos grupos nesse contexto foram bastante similares, demonstrando que a ausência do ChatGPT na prova não resultou em vantagem para os alunos do grupo experimental sobre o grupo controle.
- Nos formulários de percepção aplicados na experiência, ao decorrer das aplicações, foram introduzidas novas questões em que somente os alunos que optaram por utilizar o ChatGPT em sala de aula responderiam. Acredita-se que com essas atualizações, foi possível coletar dados mais específicos e direcionados aos estudantes que realmente utilizaram a ferramenta em sala de aula.

Capítulo 11

Limitações

Durante o andamento dos experimentos, algumas limitações foram identificadas. Estas limitações estão relacionadas ao tempo disponível para as atividades em sala de aula, questões organizacionais e também fatores externos que potencialmente impactaram os resultados obtidos.

Foi observado que em alguns dias, a maior parte dos alunos não conseguiu completar todos os exercícios propostos. Apesar de alguns alunos conseguirem concluir a atividade, outros relataram aos tutores/monitores que o tempo disponibilizado (1h20 por aula) não foi suficiente. Além disso, houve atrasos na chegada de alguns alunos à sala de aula, contribuindo para a restrição de tempo efetivo para a realização das tarefas.

No último dia de experimento, notou-se a menor taxa de presença dos alunos, especialmente do grupo controle. Esse fato pode estar relacionado com a proximidade do recesso de final de ano do mês de dezembro (Natal e Ano Novo), impactando a participação dos estudantes. Essa redução de presença pode ter interferido nos resultados comparativos entre os grupos.

Apesar dos estudantes do grupo controle serem orientados sobre a metodologia, segundo a qual não teriam permissão para acessar ferramentas de IA, como o ChatGPT, notou-se que alguns participantes apresentaram fortes indícios de uso de código gerado por IA, visto que alguns códigos dos participantes estavam idênticos aos códigos gerados por essas ferramentas pelos autores deste artigo.

Durante a correção dos exercícios enviados pelos alunos, inicialmente foi solicitado ao ChatGPT que gerasse as respostas dos exercícios para comparação com soluções apresentadas pelos alunos, na tentativa de identificar possíveis casos de plágio. Observou-se que, pelo fato de o ChatGPT gerar múltiplas versões das respostas de formas diferentes, o trabalho de verificação de plágio não poderia ser completamente preciso com o grupo de 41 alunos. Ainda assim, em alguns casos, foi possível identificar indícios de plágio, pois os códigos de alguns alunos apresentaram fortes semelhanças com os gerados pela ferra-

menta. Uma vez que não foi possível bloquear os computadores do LINF para impedir o acesso do grupo de controle a ferramentas de IA, notou-se essa limitação metodológica no processo.

Além disso, uma das limitações deste estudo está relacionado ao tamanho reduzido de participantes. Com isso, os resultados obtidos podem não capturar a diversidade de respostas ou comportamentos de uma população maior, o que limita a generalização dos resultados. Assim, futuras pesquisas são necessárias envolvendo mais alunos e turmas, para expandir os *insights* deste trabalho.

Capítulo 12

Conclusão

Neste artigo, foi apresentado um relato de experiência sobre o uso do ChatGPT em atividades de programação em sala de aula como um tutor para o ensino de Algoritmos e Programação de Computadores. Nosso trabalho avaliou o impacto do ChatGPT no desempenho acadêmico dos estudantes e analisou suas percepções sobre a ferramenta em comparação com a tutoria tradicional.

Os resultados indicam que os alunos do grupo experimental, que utilizaram o ChatGPT em sala para a resolução dos exercícios, foram capazes de entregar mais atividades em menos tempo em comparação com os alunos que não tiveram acesso à ferramenta, mesmo contando com o suporte de monitores, tutores e do docente em sala. No entanto, a qualidade dos exercícios entregues pelo grupo experimental não foi significativamente superior à do grupo controle. Dessa forma, a quantidade de resoluções apresentadas pelos alunos que utilizaram o ChatGPT não é diretamente proporcional à precisão ou qualidade das respostas fornecidas. Isso pode ser observado ao analisarmos o resultado da primeira prova, na qual ambos os grupos realizaram a avaliação da mesma maneira, sem o uso do ChatGPT ou de tutoria tradicional, e obtiveram desempenho semelhante.

Quanto à percepção do uso do ChatGPT como ferramenta de assistência, os estudantes consideraram o *chatbot* bastante útil. Entretanto, a maioria acredita que seu uso em sala de aula poderia causar um impacto negativo. Os alunos apontam que a ferramenta poderia prejudicar o desenvolvimento do raciocínio e da aprendizagem, além de levar à dependência excessiva, muitas vezes decorrente da comodidade que a ferramenta proporciona. Por último, os alunos afirmaram que ainda preferem a tutoria tradicional em comparação à tutoria baseada no ChatGPT, mas não descartam seu uso como recurso complementar no ambiente de ensino.

Como trabalhos futuros, seria interessante aplicar uma abordagem que permita verificar a maneira como os alunos estão utilizando o ChatGPT em sala de aula, de forma a investigar melhor o uso pedagógico, ou seja, avaliar se o ChatGPT oferece somente uma

resolução automática de problemas ou contribui para o desenvolvimento do raciocínio lógico dos alunos. Além disso, seria relevante a implementação de *chatbots* personalizados, que oferecem maior controle sobre as respostas, comparando seu uso e impacto no desempenho e aprendizado dos alunos com os impactos do ChatGPT nesses mesmos aspectos. Seria pertinente também adotar uma metodologia para investigar o uso do ChatGPT pelos alunos em diferentes tipos de questões, para compreender em quais tipos de atividades o assistente virtual apresenta melhor suporte. Por fim, outro ponto interessante seria avaliar o desempenho e a aceitação dos estudantes em relação ao ChatGPT em um número maior de turmas, assim como em outras disciplinas do departamento de Ciência da Computação. Isso poderia revelar possíveis percepções diferentes dependendo do nível (iniciante ou avançado) das disciplinas.

Referências

- [1] Zhai, Xiaoming: *Chatgpt user experience: Implications for education*. Preprint, 2022. <https://www.researchgate.net/publication/366463233>. 1, 3
- [2] Azaiz, Imen, Oliver Deckarm e Sven Strickroth: *Ai-enhanced auto-correction of programming exercises: How effective is gpt-3.5?* International Journal of Engineering Pedagogy (iJEP), 13:67–83, dezembro 2023. 1, 3, 4
- [3] Sun, Dan, Azzeddine Boudouaia, Chengcong Zhu e Yan Li: *Would chatgpt-facilitated programming mode impact college students' programming behaviors, performances, and perceptions? an empirical study international journal of educational technology in higher education*. International Journal of Educational Technology in Higher Education, fevereiro 2024. 1, 4
- [4] Alsobeh, Anas e Belle Woodward: *Ai as a partner in learning: A novel student-in-the-loop framework for enhanced student engagement and outcomes in higher education*. outubro 2023. 3
- [5] Marian, Marius, Ilie Borcoși, Eugen Ganea, Corina Ana Borcoși, Dorin Popescu e Adelin Cusman: *A study on the perception of digitalization and use of digital technologies in higher-education institutions*. páginas 100–105, outubro 2023. 3
- [6] Liu, Rongxin, Carter Zenke, Charlie Liu, Andrew Holmes, Patrick Thornton e David Malan: *Teaching cs50 with ai: Leveraging generative artificial intelligence in computer science education*. páginas 1927–1927, março 2024. 3
- [7] Mahmud, Arif, Afjal Sarower, Amir Sohel, Md Assaduzzaman e Touhid Bhuiyan: *Adoption of chatgpt by university students for academic purposes: Partial least square, artificial neural network, deep neural network and classification algorithms approach*. Array, 21:100339, março 2024. 3
- [8] Graven, Olaf e Lachlan Mackinnon: *Developing higher education – post-pandemic – influenced by ai*. páginas 01–08, outubro 2023. 3
- [9] Smolansky, Adele, Andrew Cram, Corina Radulescu, Sandris Zeivots, Elaine Huber e René Kizilcec: *Educator and student perspectives on the impact of generative ai on assessments in higher education*. páginas 378–382, julho 2023. 3, 4
- [10] Baidoo-Anu, David e Leticia Ansah: *Education in the era of generative artificial intelligence (ai): Understanding the potential benefits of chatgpt in promoting teaching and learning*. Journal of AI, 7, março 2023. 3, 4

- [11] Popovici, Matei Dan: *Chatgpt in the classroom. exploring its potential and limitations in a functional programming course*. International Journal of Human-Computer Interaction, 40(22):7743–7754, outubro 2023, ISSN 1532-7590. <http://dx.doi.org/10.1080/10447318.2023.2269006>. 3
- [12] Silva, Carlos Alexandre Gouvea da, Felipe Negrelle Ramos, Rafael Veiga de Moraes e Edson Leonardo dos Santos: *Chatgpt: Challenges and benefits in software programming for higher education*. Sustainability, 16(3), 2024, ISSN 2071-1050. <https://www.mdpi.com/2071-1050/16/3/1245>. 3, 4
- [13] French, Fiona, David Levi, Csaba Maczo, Aiste Simonaityte, Stefanos Triantafyllidis e Gergo Varda: *Creative use of openai in education: Case studies from game development*. Multimodal Technologies and Interaction, 7, agosto 2023. 3
- [14] Dahal, Niroj, Basanta Lamichhane, Bal Luitel e Binod Pant: *Ai chatbots as math algorithm problem solvers: A critical evaluation of its capabilities and limitations*. dezembro 2023. 3
- [15] Hajj, Jana e Melike Sah: *Assessing the impact of chatgpt in a php programming course*. páginas 1–10, novembro 2023. 3, 5, 7
- [16] Bastani, Hamsa, Osbert Bastani, Alp Sungu, Haosen Ge, Özge Kabakçı e Rei Mariman: *Generative ai can harm learning*. The Wharton School Research Paper, julho 2024. <https://dx.doi.org/10.2139/ssrn.4895486>. 4
- [17] Yilmaz, Ramazan e Fatma Gizem Karaoglan Yilmaz: *Augmented intelligence in programming learning: Examining student views on the use of chatgpt for programming learning*. Computers in Human Behavior: Artificial Humans, 1(2):100005, 2023, ISSN 2949-8821. <https://www.sciencedirect.com/science/article/pii/S2949882123000051>. 4
- [18] McLeod, Amber e Hannah Richardson: *Co-constructing skills for chatgpt at university*. Journal of Academic Language and Learning, 17(1):T70–T80, Sep. 2023. <https://journal.aall.org.au/index.php/jall/article/view/929>. 4
- [19] Gallent Torres, Cinta, A Zapata-González e José Ortego Hernando: *The impact of generative artificial intelligence in higher education: a focus on ethics and academic integrity*. RELIEVE - Revista Electrónica de Investigación y Evaluación Educativa, 29:1–19, dezembro 2023. 4
- [20] Laato, Samuli, Benedikt Morschheuser, Juho Hamari e Jari Björne: *Ai-assisted learning with chatgpt and large language models: Implications for higher education*. julho 2023. 4
- [21] Mendonça, Nabor C.: *Evaluating chatgpt-4 vision on brazil's national undergraduate computer science exam*, 2024. <https://arxiv.org/abs/2406.09671>. 5
- [22] Kaleemunnisa, FNU, Christelle Scharff, Krishna Bathula e Begimai Zhumakova: *Chatgpt in the classroom: Experimentation in a python class for non-computing majors*. Em 2024 IEEE Integrated STEM Education Conference (ISEC), páginas 1–7, 2024. 5, 6

- [23] Xue, Yuankai, Hanlin Chen, Gina R. Bai, Robert Tairas e Yu Huang: *Does chatgpt help with introductory programming? an experiment of students using chatgpt in cs1*. Em *Proceedings of the 46th International Conference on Software Engineering: Software Engineering Education and Training, ICSE-SEET '24*, página 331–341, New York, NY, USA, 2024. Association for Computing Machinery, ISBN 9798400704987. <https://doi.org/10.1145/3639474.3640076>. 5, 6
- [24] Paula, Taw Ham, Pedro Bravim, Simon Castro, Lafayette Melo e Alex Rêgo: *Aprendizagem de estrutura de dados utilizando o chatgpt como ferramenta auxiliar: um relato de experiência com o conteúdo de árvores binárias*. Em *Anais do XXXII Workshop sobre Educação em Computação*, páginas 396–406, Porto Alegre, RS, Brasil, 2024. SBC. <https://sol.sbc.org.br/index.php/wei/article/view/29643>. 6
- [25] Reynolds, Laria e Kyle McDonell: *Prompt programming for large language models: Beyond the few-shot paradigm*, 2021. <https://arxiv.org/abs/2102.07350>. 11

Apêndice A

Listas de Exercícios

Esta seção apresenta os exercícios aplicados em cada dia do experimento e seus respectivos *links* no site do *Becrowd*.

A.1 Lista de Exercícios - Dia 1

A.1.1 Exercício 1 (*Equivalente ao Becrowd 1038*)

Com base na Tabela A.1, escreva um programa em Python que leia o código de um item e a quantidade deste item. A seguir, calcule e mostre o valor da conta a pagar.

Tabela A.1: Relação de código, especificação e preço do produto

Código	Especificação	Preço
1	Cachorro Quente	R\$ 4.00
2	X-Salada	R\$ 4.50
3	X-Bacon	R\$ 5.00
4	Torrada simples	R\$ 2.00
5	Refrigerante	R\$ 1.00

Entrada: O arquivo de entrada contém dois valores inteiros correspondentes ao código e à quantidade de um item conforme Tabela A.1.

Saída: O arquivo de saída deve conter a mensagem "Total: R\$ "seguido pelo valor a ser pago, com 2 casas após o ponto decimal.

Tabela A.2: Casos de Teste do Exercício A.1.1

Exemplo de Entrada	Exemplo de Saída
3 2	Total: R\$ 10.00
4 3	Total: R\$ 6.00
2 3	Total: R\$ 13.50

A.1.2 Exercício 2 (*Equivalente ao Beecrowd 1051*)

Em um país imaginário denominado Lisarb, todos os habitantes ficam felizes em pagar seus impostos, pois sabem que nele não existem políticos corruptos e os recursos arrecadados são utilizados em benefício da população, sem qualquer desvio. A moeda deste país é o Rombus, cujo símbolo é o R\$.

Faça um programa em Python que leia um valor com duas casas decimais, equivalente ao salário de uma pessoa de Lisarb. Em seguida, calcule e mostre o valor que esta pessoa deve pagar de Imposto de Renda, segundo a Tabela A.4.

Tabela A.3: Relação entre faixa de renda e taxa de imposto de renda

Renda	Imposto de Renda
De R\$ 00.00 até R\$ 2000.00	Isento
De R\$ 2000.01 até R\$ 3000.00	8%
De R\$ 3000.01 até R\$ 4500.00	18%
Acima de R\$ 4500.00	28%

Lembre que, se o salário for R\$ 3002.00, a taxa que incide é de 8% apenas sobre R\$ 1000.00, pois a faixa de salário que fica de R\$ 0.00 até R\$ 2000.00 é isenta de Imposto de Renda. No exemplo fornecido (abaixo), a taxa é de 8% sobre R\$ 1000.00 + 18% sobre R\$ 2.00, o que resulta em R\$ 80.36 no total. O valor deve ser impresso com duas casas decimais.

Entrada: A entrada contém apenas um valor de ponto flutuante, com duas casas decimais.

Saída: Imprima o texto "R\$" seguido de um espaço e do valor total devido de Imposto de Renda, com duas casas após o ponto. Se o valor de entrada for menor ou igual a 2000, deverá ser impressa a mensagem "Isento".

Tabela A.4: Casos de Teste do Exercício A.1.2

Exemplos de Entrada	Exemplos de Saída
3002.00	R\$ 80.36
1701.12	Isento
4520.00	R\$ 355.60

A.1.3 Exercício 3 (*Equivalente ao Beecrowd 1117*)

Em Python, faça um programa que leia as notas referentes às duas avaliações de um aluno. Calcule e imprima a média semestral. Faça com que o algoritmo só aceite notas válidas (uma nota válida deve pertencer ao intervalo $[0,10]$). Cada nota deve ser validada separadamente.

Entrada: A entrada contém vários valores reais, positivos ou negativos. O programa deve ser encerrado quando forem lidas duas notas válidas.

Saída: Se uma nota inválida for lida, deve ser impressa a mensagem "nota invalida". Quando duas notas válidas forem lidas, deve ser impressa a mensagem "media = "seguido do valor do cálculo. O valor deve ser apresentado com duas casas após o ponto decimal.

Tabela A.5: Casos de Teste do Exercício A.1.3

Exemplo de Entrada	Exemplo de Saída
-3.5	nota invalida
3.5	nota invalida
11.0	media = 6.75
10.0	

A.2 Lista de Exercícios - Dia 2

A.2.1 Exercício 1 (*Equivalente ao Beecrowd 2717*)

A fabricação dos presentes para o Natal é um processo muito complicado. Diversas vezes os duendes ficam até tarde trabalhando para que tudo possa ser terminado a tempo e com perfeição.

Para melhor gerenciar seus cronogramas, os duendes estipularam quantos minutos são necessários para fabricar cada presente.

Já está quase no final do expediente, e um dos duendes pediu sua ajuda. Faltam **N** minutos para a hora de ir embora, e restam dois presentes para o duende Ed fabricar. Ajude-o a descobrir se ele conseguirá fabricar os dois ainda hoje, ou se deve deixar o trabalho para amanhã.

Entrada: Cada caso de teste inicia com um inteiro **N**, indicando quantos minutos faltam para o final do expediente ($2 \leq \mathbf{N} \leq 100$).

Em seguida haverá dois inteiros **A** e **B**, indicando quantos minutos são necessários para fabricar os dois presentes que Ed precisa fabricar ($1 \leq \mathbf{A}, \mathbf{B} \leq 100$).

Saída: Imprima uma linha, contendo a frase "Farei hoje!" caso seja possível fabricar os dois presentes antes do final do expediente, ou "Deixa para amanhã!" caso contrário.

Tabela A.6: Casos de Teste do Exercício A.2.1

Exemplos de Entrada	Exemplos de Saída
20 15 6	Deixa para amanhã!
20 10 10	Farei hoje!

A.2.2 Exercício 2 (*Equivalente ao Beecrowd 1150*)

Faça um programa que leia dois inteiros: **X** e **Z** (devem ser lidos tantos valores para **Z** quantos necessários, até que seja digitado um valor maior do que **X** para ele). Conte quantos números inteiros devem ser somados em sequência (considerando o **X** nesta soma) para que a soma ultrapasse a **Z** o mínimo possível. Escreva o valor final da contagem.

A entrada pode conter, por exemplo, os valores 21 21 15 30. Neste caso, é então assumido o valor 21 para **X** enquanto os valores 21 e 15 devem ser desconsiderados pois são menores ou iguais a **X**. Como o valor 30 está dentro da especificação (maior do que **X**) ele será válido e então deve-se processar os cálculos para apresentar na saída o valor 2, pois é a quantidade de valores somados para se produzir um valor maior do que 30 (21 + 22).

Entrada: A entrada contém somente valores inteiros, um por linha, podendo ser positivos ou negativos. O primeiro valor da entrada será o valor de **X**. A próxima linha da entrada irá conter **Z**. Se **Z** não atender a especificação do problema, ele deverá ser lido novamente, tantas vezes quantas forem necessárias.

Saída: Imprima uma linha com um número inteiro que representa a quantidade de números inteiros que devem ser somadas, de acordo com a especificação acima.

Tabela A.7: Casos de Teste do Exercício A.2.2

Exemplo de Entrada	Exemplo de Saída
3	5
1	
20	

A.2.3 Exercício 3 (*Equivalente ao Beecrowd 1146*)

Este programa deve ler uma variável inteira **X** inúmeras vezes (deve parar quando o valor no arquivo de entrada for igual a zero). Para cada valor lido imprima a sequência de 1 até **X**, com um espaço entre cada número e seu sucessor.

Entrada: O arquivo de entrada contém vários números inteiros. O último número no arquivo de entrada é 0.

Saída: Para cada número **N** do arquivo de entrada deve ser impressa uma linha de 1 até **N**, conforme o exemplo abaixo. Não deve haver espaço em branco após o último valor da linha.

Tabela A.8: Casos de Teste do Exercício A.2.3

Exemplo de Entrada	Exemplo de Saída
5	1 2 3 4 5
10	1 2 3 4 5 6 7 8 9 10
3	1 2 3
0	

A.3 Lista de Exercícios - Dia 3

A.3.1 Exercício 1 (*Equivalente ao Beecrowd 1151*)

A seguinte sequência de números 0 1 1 2 3 5 8 13 21... é conhecida como série de Fibonacci. Nessa sequência, cada número, depois dos 2 primeiros, é igual à soma dos 2 anteriores. Escreva um algoritmo que leia um inteiro N ($N < 46$) e mostre os N primeiros números dessa série.

Entrada: O arquivo de entrada contém um valor inteiro N ($0 < N < 46$).

Saída: Os valores devem ser mostrados na mesma linha, separados por um espaço em branco. Não deve haver espaço após o último valor.

Tabela A.9: Casos de Teste do Exercício A.3.1

Exemplo de Entrada	Exemplo de Saída
5	0 1 1 2 3

A.3.2 Exercício 2 (*Equivalente ao Beecrowd 3358*)

A região sul do Brasil é caracterizada pela ascendência multicultural de seus habitantes, sendo principalmente europeus e sobretudo italianos, alemães e poloneses. Uma consequência interessante disso é a variação na dificuldade na pronúncia dos sobrenomes da população, o que as vezes dificulta a vida dos professores na realização da chamada de sua turma, gerando até situações constrangedoras. Dada a possibilidade de constrangimento em suas aulas, a professora Jiraiya decidiu pesquisar os sobrenomes em sua lista de chamadas. Na concepção de Jiraiya, um sobrenome é difícil se tiver três ou mais consoantes consecutivas.

Entrada: A entrada contém vários casos de teste. A primeira linha possui um inteiro N que indica o número de casos de teste. Cada caso de teste consiste em um sobrenome. A *string* contém letras do alfabeto sem acentos, a primeira letra está sempre em maiúscula e o sobrenome pode ter no máximo 42 caracteres.

Saída: Para cada caso de entrada, imprima o sobrenome e se é fácil ou não, conforme mostra o exemplo abaixo.

Tabela A.10: Casos de Teste do Exercício A.3.2

Exemplo de Entrada	Exemplo de Saída
6	Ferrari eh facil
Ferrari	Bianchi nao eh facil
Bianchi	Hoffmann nao eh facil
Hoffmann	Hofmann eh facil
Hofmann	Lewandowski nao eh facil
Lewandowski	Nowak eh facil
Nowak	

A.3.3 Exercício 3 (*Equivalente ao Beecrowd 2137*)

Em uma pacata cidade do interior, o senhor Severino decidiu montar a própria biblioteca, já que coleciona vários livros desde sua juventude. Como ele não sabe programar, pediu ajuda ao neto para criar um programa que cadastre e ordene seus livros pelo código. Porém, seu neto ainda está no ensino fundamental, e como sabe muito pouco de programação, acabou criando um programa que somente cadastra os livros, mas não os ordena. Desse modo, o senhor Severino recorreu a você, pois sabe de suas habilidades com programação. Sua tarefa é simples: ordenar os cadastros dos códigos dos livros.

Entrada: A entrada contém vários casos de teste. Cada teste começa com um valor N ($1 \leq N \leq 1000$). Em seguida, N linhas terão os códigos dos livros, que estão sempre no formato “xxxx”, isto é, não haverá o cadastro ‘1’, por exemplo, mas “0001”. A entrada termina com a linha com um único “0”.

Saída: Seu programa deverá imprimir o cadastro dos códigos ordenado. Não haverá linha em branco entre os casos de teste.

Tabela A.11: Casos de Teste do Exercício A.3.3.

Exemplo de Entrada	Exemplo de Saída
6	0015
3	0100
1233	1233
0015	0000
0100	0001
7	0752
0752	1110
1110	6321
0001	6322
6322	8000
8000	
6321	
0000	
0	

Apêndice B

Listas de *Prompts*

Esta seção apresenta os *prompts* fornecidos como apoio para os alunos do grupo experimental. Entre cerquilhas estão os elementos em que os alunos deveriam preencher com o texto do exercício, dúvidas, código, ou casos de teste.

B.1 Para dúvidas com o Exercício

B.1.1 *Prompt 1*

Poderia me explicar o que devo fazer nesse exercício? Segue o enunciado abaixo:
[EXERCÍCIO]

B.1.2 *Prompt 2*

Estou com dúvida sobre [DÚVIDA] no exercício abaixo, poderia me esclarecer?
[EXERCÍCIO]

B.1.3 *Prompt 3*

Estou com dúvida em como implementar o seguinte exercício: [EXERCÍCIO]
Poderia me dar dicas sobre como implementar?

B.2 Para depuração de código

B.2.1 *Prompt 4*

Meu código em Python está dando erros, este deveria reproduzir [EXPLIQUE O QUE O CÓDIGO DEVERIA FAZER]. Qual poderia ser o erro?

O código com erro é: *[SEU CÓDIGO]*

Está gerando a seguinte saída: *[SAÍDA DO SEU CÓDIGO]*

A saída esperada seria: *[SAÍDA ESPERADA]*

Por favor, somente indique o erro.

B.3 Ajuda com ideias e melhorias de código

B.3.1 *Prompt 5*

Estou com dúvidas sobre qual seria a abordagem correta para implementar em Python *[O QUE VOCÊ QUER FAZER]*. Por favor, me indique maneiras de fazer isso.

B.3.2 *Prompt 6*

Como eu poderia melhorar o seguinte código em Python *[CÓDIGO]*? Poderia me passar dicas e apontar erros no código que poderiam estar impedindo a melhor performance?