



On the Characterization of Behavior-Driven Development Adoption Benefits: A Multiple Case Study

Thiciane Couto
PPGCC - Escola Politécnica - PUCRS
Porto Alegre, RS, Brazil
thiciane.silva@edu.pucrs.br

Sabrina Marczak
PPGCC - Escola Politécnica - PUCRS
Porto Alegre, RS, Brazil
sabrina.marczak@pucrs.br

Daniel Callegari
PPGCC - Escola Politécnica - PUCRS
Porto Alegre, RS, Brazil
daniel.callegari@pucrs.br

Michael Móra
PPGCC - Escola Politécnica - PUCRS
Porto Alegre, RS, Brazil
michael.mora@pucrs.br

Fabio Gomes
Instituto de Tecnologia e Pesquisa
Universidade Tiradentes - UNIT
Aracaju, SE, Brazil
fabio_gomes@unit.br

RESUMO

BDD supports software development teams in building and delivering user needs. Although the claims that BDD adoption improves product and software processes, there is still no consolidated evidence about its benefits nor have we found quality indicators to measure them. This paper reports on a study that aimed to identify and characterize the benefits of BDD adoption in software teams novice to it. We conducted a multiple case study in teams formed by students working with industry practitioners in active-based learning programs. Perceived benefits are improved communication, faster delivery cycles, better product quality, among others. Each benefit was characterized into aspects that express how the benefit is conceptualized. For instance, improved communication is perceived by sharper conversations with the product owner. Our results might inspire both practitioners and scholars in the pursuit of quality indicators to measure the benefits of adopting BDD.

CCS CONCEPTS

• Software and its engineering; • Software development methods;

KEYWORDS

Behavior-Driven Development, BDD Adoption, Benefits, Case Study

ACM Reference Format:

Thiciane Couto, Sabrina Marczak, Daniel Callegari, Michael Móra, and Fabio Gomes. 2022. On the Characterization of Behavior-Driven Development Adoption Benefits: A Multiple Case Study. In *XXI Brazilian Symposium on Software Quality (SBQS '22)*, November 7–10, 2022, Curitiba, Brazil. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3571473.3571492>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

SBQS '22, November 7–10, 2022, Curitiba, Brazil

© 2022 Association for Computing Machinery.

ACM ISBN 978-1-4503-9999-9/22/11...\$15.00

<https://doi.org/10.1145/3571473.3571492>

1 INTRODUÇÃO

O *Test-Driven Development* (TDD) foi originado como prática do *Extreme Programming* (XP) e passou a ser utilizado como apoio à realização de testes unitários no desenvolvimento de software [11]. Com a evolução do seu uso percebeu-se que emergiam dúvidas sobre os requisitos que são base para a especificação dos cenários de teste. Assim, Dan North [1] idealizou o *Behavior-Driven Development* (BDD), cujo objetivo é descrever o comportamento do sistema e, assim, proporcionar uma melhor comunicação entre os envolvidos no desenvolvimento de software e, conseqüentemente, a redução dos problemas ocorridos com o TDD.

Smart [21] argumenta que a adoção do BDD pode proporcionar benefícios tais como mudanças seguras, *releases* mais rápidas, melhoria na comunicação, redução de desperdício e de custo. Ainda são escassas as evidências de que estes benefícios de fato ocorrem, em especial em equipes novas na sua adoção.

Visto que Kamalakar *et al.* [9] apontam o crescimento emergente da adoção desta abordagem por profissionais da indústria, buscou-se evidências junto a este público-alvo a partir de uma revisão da *grey literature* [7]. Identificou-se nesta revisão [5] o apontamento de benefícios tais como melhoria da comunicação e da qualidade, entretanto, sem detalhes sobre como estes benefícios foram medidos e em qual contexto ocorreram. Estes resultados ecoaram os exíguos resultados de uma revisão exploratória da literatura científica anterior, brevemente também registrada em [5]—ausência de indicadores e benefícios listados sem riqueza de detalhes de seus significados. Por exemplo, Irshad, Britto e Petersen [8] afirmam que a adoção do BDD em equipes de larga escala facilitou a reutilização de artefatos e auxiliou na organização dos testes sem apontar a extensão dos benefícios ou formas de mensurá-los.

Dada essa lacuna na literatura, definiu-se como objetivo de pesquisa, à longo prazo, a definição de indicadores de qualidade da adoção do BDD em equipes de desenvolvimento de software. Antes de chegar-se aos indicadores, é preciso, entre outros, melhor entender como os benefícios se caracterizam em si. Por exemplo, para poder-se definir um indicador para o benefício da melhoria da comunicação, é preciso entender como a comunicação propriamente dita melhorou: reduziu-se o número de trocas de mensagens? o número de reuniões com o cliente? Ou seria uma maior riqueza de detalhes que podem ser traduzidos nas cláusulas que compõem os cenários BDD de uma *feature*? Assim, *caracterizar os benefícios*



Figura 1: Processo do BDD (Adaptação de: [21])

percebidos por equipes novatas na adoção do BDD é o objetivo do estudo de caso múltiplo apresentado neste artigo—o primeiro passo rumo ao objetivo geral desta pesquisa.

A partir deste estudo de caso múltiplo, este artigo traz três principais contribuições: (i) um conjunto de benefícios percebidos pelos integrantes das equipes, (ii) os aspectos que caracterizam estes benefícios no contexto dos casos estudados, e (iii) evidências empíricas para adicionar ao corpo de conhecimento do tópico. O restante do artigo discorre sobre o estudo e as contribuições apontadas.

2 BEHAVIOR-DRIVEN DEVELOPMENT

O BDD é flexível e pode ser utilizado em todo o ciclo de vida de desenvolvimento de software ou em apenas uma parte dele, como no apoio à especificação dos requisitos [21]. Assim, ele pode auxiliar no entendimento dos interesses do cliente pela equipe de desenvolvimento. Por isso, pode-se destacar que a grande contribuição do BDD é a integração de todos os envolvidos no projeto, onde eles possuem acesso a qualquer aspecto do software (requisito, código, testes). De acordo com Solis e Wang [22], o BDD é a combinação das seguintes características e práticas ágeis:

- Linguagem ubíqua: Semelhante à linguagem comum, com termos de fácil entendimento. Qualquer membro da equipe pode entender e discutir o que está sendo desenvolvido.
- Testes de aceitação automatizados: Após o entendimento dos critérios de aceitação, é tarefa da equipe definir testes (das funcionalidades) de software a partir desses critérios.
- Escrita legível do código: A implementação deve ser clara e expressar o objetivo dos cenários BDD. No momento da implementação, o código (métodos, classes, etc.) deve indicar o seu objetivo de acordo com as especificações que foram disponibilizadas nas *user stories*.

O BDD também visa promover a especificação dos requisitos por meio de exemplos. Apesar disso, seu objetivo não é a especificação de critérios de aceitação, automação e nem testes, embora ele contenha esses elementos [22]. Essa especificação auxilia na conexão do código com os requisitos, gerando assim, um ambiente de gerenciamento do progresso do desenvolvimento e dos requisitos durante todo o ciclo de desenvolvimento [21]. Ainda, conforme Smart [21], a evolução do BDD ampliou a visão da análise ágil e, em consequência, no auxílio da entrega de valor.

Cenário [01]: Como usuário desejo atualizar minha senha

Dado que estou na página de perfil do Admin
Quando eu atualizo a senha
Então recebo uma mensagem de senha atualizada com sucesso
E a senha é atualizada

Figura 2: Exemplo do uso da linguagem Gherkin

Com essa evolução emergiu um processo composto por seis passos interativos e que podem ser aplicados junto às metodologias ágeis [20], conforme ilustrado na Figura 1. O ciclo é iniciado com a fase de planejamento para definição das expectativas de negócio. Na etapa de análise, os objetivos de negócio são decompostos em *features*, que são associadas a uma ou mais *user stories*. Essas *user stories* são descritas em múltiplos cenários BDD, onde cada cenário possui um comportamento de acordo com a situação. Esses cenários são estruturados segundo a linguagem ubíqua Gherkin [22] (vide exemplo na Figura 2)—Dado que “contexto”, Quando “evento”, Então “resultado”. O operador lógico “E” é utilizado para adicionar condições. Para finalizar esses passos, são realizadas entregas ao cliente, que é quem faz a validação do produto entregue, reiniciando o ciclo ao discutir novas *features* ou ao solicitar correções.

Dado que os cenários BDD são a base para as atividades do ciclo, Oliveira e Marczak [13] argumentam que é importante que esses cenários apresentem uma boa qualidade. Cenários com escopo inconsistente ou incompletos devem gerar problemas de comunicação entre os envolvidos e introduzir defeitos no código a ser entregue. Esses autores junto a Moralles [14] propuseram um *checklist* para avaliar a qualidade de um cenário BDD (apesar do argumento ainda não ter sido validado), que considera os seguintes atributos de qualidade: conciso, viável, testável, compreensível e não ambíguo.

Esse contexto inspira o interesse em melhor entender como se caracterizam os benefícios percebidos da adoção do BDD. Em especial, quando há estudos prévios como o de Kamalakar et al. [9] que argumentam que são limitados os estudos científicos que discorrem sobre o uso do BDD e seus benefícios na prática. Assim, como previamente mencionado, para preencher essa lacuna, o objetivo desse artigo é relatar um estudo de caso múltiplo que visou caracterizar os benefícios identificados na adoção do BDD sob a perspectiva de equipes novatas a esta abordagem.

3 TRABALHOS RELACIONADOS

Binamungu, Embury e Konstantinou [4] investigaram o uso do BDD, seus benefícios e os desafios de manter as especificações do BDD na prática. A partir da realização de entrevistas com 75 profissionais de 26 países, os autores identificaram que o uso de termos específicos de domínio promovem melhorias na comunicação entre as partes interessadas. Dessa forma, ocorre a facilitação da compreensão dos objetivos principais das necessidades de usuário definidas. Assim, os autores concluíram que o uso do BDD promove benefícios como: melhoria na comunicação, facilidade de entendimento das *features* e, consequentemente, entregas mais rápidas do que é solicitado.

Rocha et al. [17] apresentam a aplicação do uso conjunto do TDD e BDD no ensino de Engenharia de Software. Para isso, os autores

realizaram uma investigação empírica, em uma turma do curso de Bacharelado em Tecnologia da Informação do último nível do curso em uma universidade privada. Essa experiência consistiu em um estudo de caso com estudantes que foram divididos em grupos e tinham como objetivo a entrega de um software que foi determinado pelo professor. Como resultados, os autores relataram que nesta disciplina houve uma redução nas ausências dos estudantes, uma maior taxa de satisfação dos estudantes, notas mais elevadas, um aprendizado na implementação de teste devido à utilização do TDD e à adoção do BDD, como forma de melhorar a comunicação, entre-gas e diminuir o retrabalho. Além disso, essa abordagem permitiu que os estudantes entregassem um produto em um curto tempo.

O estudo de Nascimento *et al.* [12] também apresenta um estudo de caso, porém diferentemente do estudo anterior [17], esse estudo relata o impacto do BDD em equipes de desenvolvimento ágil de software. O estudo de caso foi realizado em um ambiente de desenvolvimento de aplicativos *mobile* e contou com 42 participantes. Os autores realizaram entrevistas com esses participantes. Resultados sugerem que a adoção do BDD gerou impactos positivos como o aumento da colaboração entre os membros da equipe, redução de ambiguidades, melhoria na divisão das tarefas, na organização e entendimento do projeto. Porém, os autores também citam impactos negativos como a dificuldade na escrita dos testes e na implantação do TDD. Apesar das dificuldades, os participantes do estudo manifestaram perceber que os aspectos positivos superaram os negativos.

Os trabalhos aqui citados—[4], [17] e [12]—relatam a respeito da adoção e os benefícios da adoção do BDD por profissionais e estudantes. Estes trabalhos apresentam como resultados benefícios após a adoção do BDD, como o aumento da colaboração, melhorias na comunicação, na organização, redução no retrabalho e no tempo de entrega. Porém, estes resultados não discorrem sobre como esses benefícios são caracterizados. Também, nenhum dos estudos considera equipes novatas no uso do BDD, seja na indústria ou não, oportunizando o estudo de caso múltiplo relatado neste artigo.

4 ESTUDO DE CASO MÚLTIPLO

4.1 Estudo de Caso Tradicional e Participativo

Essa pesquisa investigou o caso de equipes novatas na adoção do BDD. Entende-se por novata equipes que estavam sendo expostas pela primeira vez ao uso do BDD na prática, tendo ou não conhecimento teórico prévio. Por conveniência, estas equipes são constituídas de estudantes de graduação em formação de cursos complementares fomentados e em atuação com empresas. Representantes das empresas desempenham o papel de cliente e usuário e atuam junto às equipes de desenvolvimento, seguindo a filosofia ágil. O processo de capacitação dos estudantes é mediado por professores, também especialistas em BDD, que atuam como mentores (ou, se preferir, *agile coach*, para usar a nomenclatura ágil).

Também por conveniência, utilizou-se duas categorias de estudo de caso—o “Tradicional” [18] e o Estudo de Caso Participativo [16], para refletir a natureza das relações dos pesquisadores com as empresas que fomentam os cursos de capacitação em que os casos foram investigados. Em duas delas, Caso 1 e Caso 2 (detalhes a seguir), os pesquisadores já atuavam diretamente como integrantes da equipe do curso de capacitação, conduzindo então um Estudo de Caso Participativo, onde é permitido ao pesquisador ser ativo na interação com o caso observado e ter atuação junto a ele. Já no

Caso 3, teve-se acesso como ouvinte, a partir do contato do último autor, caracterizando o Estudo de Caso Tradicional.

Cabe destacar que um Estudo de Caso Participativo difere-se da Pesquisa-Ação [2] pois essa implica que a atuação do pesquisador irá incidir na transformação da realidade no ambiente de estudo. Ou seja, o pesquisador intencionalmente promove uma ação no objeto de estudo para observar as consequências desta ação [3]. Os casos observados já estavam usando o BDD em suas turmas de estudantes anteriores, não sendo novidade para os programas de capacitação em si essa adoção. Ou seja, os estudantes já iniciam suas atividades adotando desde o início do processo de desenvolvimento o BDD.

Apesar da diferença do grau de interação do pesquisador entre as duas categorias de estudo de caso, Williams e Keady [24] apresentam um conjunto de valores que são comuns a esses dois tipos de estudo e que devem ser aplicados pelos pesquisadores durante sua realização. Estes valores foram inspirados nas seguintes diretrizes: i) criação de condições para o engajamento dos participantes, ii) relação entre os participantes e o estudo de caso, e iii) a existência de um espaço que possa estimular a criatividade. A partir da interconexão dessas diretrizes, têm-se os valores a seguir:

- *Analítico*: É baseado na análise do contexto do estudo de caso participativo. Por meio desse valor, o pesquisador constrói um relacionamento com os participantes, visando o engajamento deles.
- *Democrático*: Têm como base a negociação do pesquisador com os participantes a respeito da divulgação dos resultados e experiências do estudo de caso.
- *Engajamento*: Visa a garantir a facilidade e acessibilidade da aplicação do estudo de caso. E assim, engajar os participantes durante a realização do estudo.
- *Empatia*: Voltado para o pesquisador. Esse deve demonstrar sensibilidade, compassividade e positividade nos momentos em que os participantes expressarem suas opiniões e experiências, independente de sua opinião e valores pessoais.

Assim, realizou-se o estudo de caso múltiplo de acordo com esses valores como segue: os autores consideraram que o fenômeno sob investigação busca entender a perspectiva da equipe, com base no valor *analítico* e de *empatia*. Além disso, antes da realização do estudo, os autores enviaram um termo de consentimento para os participantes conforme o valor *democrático*. Durante a realização do estudo, os autores buscaram manter o *engajamento* de todos os participantes. Assim, todos os valores foram aplicados durante o planejamento e execução dessa pesquisa.

4.2 Descrição dos Casos

Como mencionado, esse estudo de caso múltiplo investigou três casos. A Tabela 1 resume as características de cada caso. Cabe ressaltar que os casos foram acompanhados de maneira virtual dada a pandemia do Covid-19.

Caso 1 - Participativo. É uma parceria da PUCRS de mais de 20 anos com uma multinacional da área de TI, denominada ORG por questões confidenciais. Anualmente, 45 estudantes, organizados em três equipes de 15 integrantes, são formados nesse curso complementar em parceria com a ORG que visa oferecer a estudantes de graduação de cursos de computação a oportunidade de trabalhar em equipe e desenvolver competências diversas, em especial as denominadas *soft skills*, como comunicação, gestão de conflitos,

Tabela 1: Caracterização dos Casos Investigados

Características	Caso 1 - Participativo	Caso 2 - Participativo	Caso 3 - Tradicional
Parceira	ORG, empresa multinacional da área de TI	Órgãos públicos de certo estado do Brasil	Órgãos Públicos e Privados do mesmo estado
Nro. de estudantes	45 divididos em 3 equipes	47 estudantes	6 estudantes
Linguagens de programação e banco de dados	C#, Web Service, Flutter, Dart, Aurélio, Banco de dados em SQL	Java, React (JavaScript), Python, Banco de dados em PostgreSQL, MySQL e MongoDB	Python com Flask, Banco de dados em PostgreSQL, MySQL e MongoDB
Frameworks/ferramentas	Azure, VS Code, Android Studio, Microsoft Teams, Zoom	Cucumber, JBehave, OpenCV, JavaCV, Tesseract, React, Node.js, Github, Trello, Discord	Behave, Trello, Discord
Utiliza versionamento?	Sim	Sim	Sim
Metodologias ágeis	Scrum, Kanban, XP	Scrum e Kanban	Scrum e Kanban
Possui capacitação dos estudantes?	Sim, antes do projeto <i>hands-on</i> , eles recebem treinamento para nivelamento e em tecnologias	Sim, incluindo BDD, OCR e os processos dos órgãos públicos	Sim, em Python, BDD, TDD e DDD
Responsáveis pela escrita das <i>user stories</i> e critérios de aceitação	PO e analistas de negócio (colaboradores)	Equipe de desenvolvimento e coordenador do projeto	PO e analistas de negócio
Adoção do BDD	Especificação das <i>user stories</i>	Ciclo completo	Ciclo completo

confiança. O programa de capacitação tem se realizado no modelo atual desde 2016, com a introdução do uso de BDD na primeira turma de 2018. O modelo envolve: um mês de capacitação teórica técnica para nivelamento de conhecimento (e.g., programação, teste e arquitetura de software, tecnologias), três meses de desenvolvimento do produto de maneira *hands-on* com o cliente e usuários (da empresa) e mentor (professor), e até 20 meses de estágio na ORG, conforme disponibilidade dos estudantes e interesse da ORG em, eventualmente, contratá-lo. O produto em desenvolvimento é do portfólio da empresa e os estudantes possuem a responsabilidade de entregar um protótipo funcional durante a etapa do *hands-on* dentro dos padrões de qualidade determinados pela empresa. Os colaboradores atuam como clientes e usuários finais do produto de software, representando o papel de *Product Owner* (PO) de uma equipe ágil. A equipe adota práticas do método ágil Scrum acompanhadas de outras práticas como Kanban e Programação em Pares, do Extreme Programming (XP). A autora que conduziu diretamente a coleta e análise de dados é aluna de pós-graduação com experiência prévia na indústria e atua 25 horas por semana no papel de apoio do mentor, auxiliando os estudantes no dia a dia. O BDD é adotado na especificação das *user stories* (US), para apoiar a definição dos critérios de aceitação que são descritos com a linguagem ubíqua Gherkin. Outros dois autores atuam na coordenação da parceria.

Caso 2 - Participativo. Parceira entre a UNIT, onde a primeira autora cursou seu bacharelado, e órgãos públicos de Sergipe. De forma análoga, os dois projetos desenvolvidos contam com a participação de 47 estudantes de bacharelado e profissionais da indústria. O projeto tem como meta digitalizar documentos públicos (diferentes clientes) e promover a redução de custo com a utilização de papéis, eliminando riscos de extravio ou perda por desgaste e maior segurança no sigilo das informações. Esse processo de digitalização envolve técnicas de tipificação, organização e limpeza dos documentos. Além disso, o projeto também implementa um sistema de gestão dos documentos a fim de facilitar a busca e identificação

deles pelos servidores. Por isso, é desenvolvida uma ferramenta de processamento e análise de conteúdo, o sistema de armazenamento e a interface de acesso do usuário. A adoção do BDD iniciou-se junto ao início do projeto e é usado em todo o ciclo de desenvolvimento, conforme ilustrado previamente na Figura 1. A equipe de desenvolvimento, composta pelos estudantes, professor coordenador, pesquisador e clientes, é responsável pela escrita das *user stories* e dos critérios de aceitação. As equipes também utilizam Scrum e Kanban. A primeira autora, na qualidade de ex-estudante do grupo de pesquisa e pesquisadora convidada, em suas 15 horas de alocação semanal, tem como responsabilidade documentar o histórico de desenvolvimento do projeto de forma a prover à equipe um relato que lhe permita, periodicamente, revisitar os processos e melhorar suas práticas. As discussões são conduzidas por ela.

Caso 3 - Tradicional. Também formado pela parceria entre a UNIT e outros órgãos públicos do Sergipe. O único projeto desse caso, quando da realização do estudo, visa desenvolver uma solução que auxilie gestores públicos e privados do estado nos processos de tomada de decisão. Conta com a participação de 8 participantes: seis deles são estudantes de graduação da área de TI, divididos em equipes de *backend* e *frontend*, um PO e um gestor público responsável pelas demandas do projeto. O PO e o gestor possuem experiência prévia com desenvolvimento de software na indústria e usando BDD e os estudantes passam inicialmente por um treinamento teórico/prático no assunto bem como sobre métodos ágeis e programação para poderem atuar no projeto. O BDD é usado junto ao Scrum e ao *Domain-Driven Development* (DDD), adotado para a definição da arquitetura do produto. De forma semelhante ao Caso 2, o BDD é utilizado em todo o ciclo, onde o PO em conjunto com os analistas de negócio dos órgãos governamentais são responsáveis pela criação das *user stories* e critérios de aceitação. Posteriormente, estes critérios são utilizados pela equipe para automatizar testes, iniciar o desenvolvimento, realizar as entregas e, assim, o ciclo é novamente iniciado até a conclusão do produto. Diferente dos

Tabela 2: Roteiro de perguntas para o *focus group*

ID	Questão
Q1	Como a equipe descreve o produto em que estão trabalhando atualmente?
Q2	Como o BDD está sendo adotado no desenvolvimento deste produto?
Q3	Quais ferramentas/ <i>frameworks</i> a equipe utilizou no auxílio da adoção do BDD?
Q4	Quais as dificuldades que a equipe passou com o uso da adoção do BDD?
Q5	Quais são os benefícios da adoção do BDD e como esses são percebidos durante o dia a dia da equipe?

outros projetos, neste o primeiro autor participou apenas como pesquisador, coletando dados para o estudo apenas.

4.3 Coleta de Dados

O principal método de coleta foi o *focus group* visando promover um momento de reflexão em grupo e facilitar o contato concomitante com os integrantes dada a virtualidade imposta pela pandemia. Também se usou da observação como recurso complementar.

Focus Group. Optou-se por um *focus group* [6] para promover o debate entre os estudantes. Utilizou-se as perguntas apresentadas na Tabela 2 para buscar identificar como os benefícios do BDD são caracterizados. Estas perguntas foram validadas por um profissional que utiliza o BDD há mais de cinco anos e por um especialista deste tipo de método de coleta. Note que as perguntas 1 a 3 são contextuais, definidas como estratégia “quebra-gelo”, para motivar os estudantes a conversarem e debaterem respeitosamente suas opiniões. A pergunta 4 foi sugerida por um dos revisores do *script* como forma de poder instigar os participantes a uma análise crítica antes de se chegar no foco do estudo, expressado na pergunta 5. Ainda, de forma geral, os participantes foram encorajados a não julgar os comentários dos colegas, respeitando a opinião de todos, para que se pudesse chegar no debate do assunto.

Foram realizadas seis sessões de coleta de dados, mediadas pelo primeiro autor, como segue: uma sessão com cada turma do Caso 1, totalizando 3 sessões; duas sessões com os estudantes do Caso 2, organizados em 2 grupos dado o alto número de estudantes deste caso; e uma sessão com o Caso 3. As sessões ocorreram após a última *sprint* de desenvolvimento, quando o produto havia sido entregue ao cliente, por completo. Cada sessão durou, em média, 90 minutos e foi organizada da seguinte forma: recrutamento e realização, que envolveu apresentar o objetivo da atividade, conduzir o debate entre os estudantes, e oferecer uma síntese para concluir a sessão. As sessões foram gravadas e transcritas, conforme Seção 4.4.

Observação. Para os Casos 1 e 2, essa observação foi participante, conforme definido por Vogt [23]. Para Vogt, o pesquisador também participa da equipe em que ele está estudando durante a coleta de dados. Malinowski [10] defende que esse tipo de participação visa reduzir a estranheza recíproca, levando o pesquisador a compartilhar de hábitos dos grupos observados. O primeiro autor organizou um diário de bordo para cada projeto de forma a registrar suas percepções sobre essas sessões. As anotações nesse diário foram guiadas por três itens em especial: o que foi observado sobre o uso do BDD, como benefícios estão sendo percebidos pelas equipes e as

Equipe	Equipe C
Data	17/11/2021
Principais fatos do dia sobre o BDD	Surgiram dúvidas de como o BDD era adotado pela equipe e o participante P11 ¹ que já havia utilizado BDD em outro emprego explicou aos demais, como ele era adotado e que neste projeto ele estava sendo adotado nos critérios de aceitação. Após essa explicação os participantes P9 e P12 foram observar os critérios das US
Benefícios do BDD	Durante a observação o participante P12 afirmou que apesar de não saber que aquilo fazia parte do BDD estava ajudando ele a entender como deve ser o comportamento e funcionamento da US que ele deveria desenvolver. O participante P9 comentou "... desenvolver com os critérios de aceitação descritos de forma simples é mais fácil, antes de colocar algo em Done é só vir aqui e olhar se está tudo ok."
Dificuldades do BDD	De volta ao grupo com os participantes comentaram que o difícil foi saber o nome do que eles estavam usando.
Características dos benefícios	O participante P11 falou: "No meu emprego passado quando a gnt usou o BDD melhorou a comunicação. O pessoal conseguia entender melhor as US, as tarefas voltavam menos também, porque verificávamos sempre os critérios"

Figura 3: Exemplo do Ciclo 1 de Análise

Código	Evidências
Reações/emoções (Emotion Coding)	P9 e P12 foram observar os critérios das US
	No início fiquei em dúvida de como utilizar.
	P20 comentou que deseja trabalhar com BDD no futuro
Palavras dos participantes (In-Vivo Coding)	difícil foi saber o nome do que eles estavam usando
	desenvolver com os critérios de aceitação descritos de forma simples é mais fácil
	quando a gnt usou o BDD melhorou a comunicação
	entender melhor as US
	a dinâmica da equipe ficou mais fácil, não tivemos um grande numero de retrabalho
foi complicado entender várias US, mas depois foi tranquilo	
com o uso do BDD fazemos os testes antes e isso facilita entendimento do TDD	
verificávamos sempre os critérios	

Figura 4: Exemplo do Ciclo 2 de Análise

dificuldades da adoção do BDD. Já para o Caso 3, as observações corresponderam a algumas sessões previamente agendadas com a equipe onde atividades em grupo estavam ocorrendo, como, por exemplo, alguma reunião com o cliente para especificação das *user stories*. Todas as anotações, dos três casos, foram posteriormente analisadas visando complementar e enriquecer os resultados das sessões de *focus group*. Desta forma, os *insights* advindos destes dados estão embutidos e dissolvidos na narrativa dos resultados.

4.4 Análise dos Dados

Utilizou-se o método de análise proposto por Saldaña [19], que é dividido em dois ciclos de codificação como apresentado a seguir. Esse processo foi apoiado pela ferramenta Atlas.TI¹, versão 7.

Ciclo 1. Tanto as transcrições dos áudios como as anotações foram analisadas em busca de dois tipos de informação: aquelas relacionadas aos sentimentos e reações dos participantes ao ouvirem o termo BDD ser mencionado pelo pesquisador ou por seus colegas, identificados com a *tag Emotion Coding* e na cor rosa; e aquelas relacionadas a palavra dita em si pelos participantes, marcada com a *tag In-Vivo Coding* na cor verde. Ambas categorias foram escolhidas dentre as propostas por Saldaña [19] por representar a natureza dos dados coletados. Segundo Patton [15], cada pesquisa é única e deve utilizar as categorias de código que lhe forem próprias. Saldaña [19] apenas ressalta que ao escolher as categorias do código estas não devem se sobrepor, como é o caso das categorias adotadas nessa pesquisa. Adotaram-se cores diferentes para grafar os trechos codificados para facilitar a análise como ilustrado na Figura 3 e Figura 4 para os Ciclos 1 e 2, respectivamente.

Ciclo 2. Após a classificação inicial, identificou-se os códigos em si a partir dos trechos separados. Por exemplo, a evidência “quando a gnt usou o BDD melhorou a comunicação” gerou o código “comunicar

¹<https://atlasti.com/>

de forma mais clara”; grafado de forma sublinhada na Seção 5.4. A seguir, trechos codificados que tinham similar conotação foram agrupados e, eventualmente, o código revisitado. Estes códigos foram, por sua vez, também agrupados, formando as categorias de resultados. A narrativa dos resultados é apresentada a partir destas categorias; grafadas no texto da Seção 5.4 em negrito. Estas mesmas categorias estão listadas na Tabela 3 (coluna “Benefício”, Seção 7).

4.5 Limitações do Estudo

Esse estudo possui um conjunto de limitações (ou ameaças à validade, conforme termo definido por [18]), quais sejam: os casos observados serem constituídos de estudantes. Dado que almejava-se estudar equipes novatas no uso do BDD, por conveniência optou-se por estudar equipes engajadas em cursos de capacitação em parceria com empresas apresentadas visto o cenário inesperado que se instalou com a pandemia. Chegou-se a negociar acesso a duas equipes da empresa ORG mas o mesmo não foi efetivado dado o foco dos colaboradores em se organizarem para mudar a operação para o virtual. Apesar disso, o cenário selecionado permitiu a possibilidade de seguir-se atuando junto à indústria: as equipes tinham como cliente e usuários os colaboradores das empresas e precisavam entregar um produto para elas desenvolvido conforme os processos das instituições. Assim, acredita-se não apenas que essa limitação é aceitável como foi bem-vinda. Os estudantes são futuros profissionais e teve-se a oportunidade de convidá-los a refletir sobre a adoção de uma prática que lhes era desconhecida.

O fato de ter-se investigado seis equipes em três casos distintos busca tratar a limitação de generalização de resultados. Ainda não se pode argumentar que os resultados são para todos os perfis de caso, mas dada a repetição entre as equipes, pode-se dizer que se deu um passo em busca dessa generalização no referido contexto.

Outra limitação imposta pela pandemia do Covid-19 foi o fato de não se poder estar em tempo integral com os estudantes de maneira face a face. As atividades da pesquisa ocorreram durante o horário dos programas de capacitação para os Casos 1 e 2, participativo, permitindo ao pesquisador interagir constantemente com os integrantes das equipes. Entretanto, como se havia atuado em 2019 de forma presencial nesses Casos, acredita-se que um pouco da riqueza da interação foi perdida quando integrantes fechavam suas câmeras, ficavam em silêncio, etc., reações observadas em outros cenários, como em sala de aula, e até mesmo por integrantes da indústria. Mitigou-se essa limitação sempre tentando engajar o estudante em conversas informais ou com atividades quebra-gelo.

Outra limitação geralmente apontada por sessões de observação é o fato da falta de riqueza pelo tempo comumente limitado de interação com os participantes. Acredita-se que, conforme argumentado por Vogt [23], o fato do estudo ser participativo permitiu uma aproximação com os participantes do estudo. Essa aproximação, por sua vez, pode ser interpretada como introdução de *viés* na interpretação dos resultados. Para mitigar essa possibilidade, corroborou-se frequentemente as anotações do principal pesquisador com os demais integrantes da equipe de pesquisa, atuando também junto aos casos, em diferentes graus de envolvimento. A discussão e alinhamento dos entendimentos serviu como contingência para possível *viés*.

Por fim, a coleta de dados foi baseada em dados providos pelos participantes do estudo. Por questões contratuais, tinha-se acesso

ao ambiente de desenvolvimento e aos artefatos como as *user stories* e os cenários BDD, mas não a *dashboards* de acompanhamento de métricas ou similares para se poder observar qualquer relação entre as melhorias percebidas e as possíveis influências dessas melhorias nos artefatos e processos. Almeja-se tratar essa limitação em um estudo futuro, negociando acesso a estas informações em novas turmas a serem estudadas para estes casos.

5 RESULTADOS

Para os casos investigados, todos os integrantes de cada uma das equipes aceitaram participar da nossa pesquisa, independente do seu papel (estudantes, professor ou colaborador das empresas). Entretanto, para o *focus group*, apenas estudantes, voluntariamente, aceitaram nosso convite, como segue: Caso 1, 12 estudantes (dos 45, 26.6%) distribuídos da seguinte forma - 4 estudantes, denominados Participante 1 (P1), P2, P3 e P4, da primeira turma, denominada Equipe A; 4 da Equipe B (P5–P8); e 4 da Equipe C (P9–P12). Para o Caso 2, participaram 8 estudantes (dos 47, 17%), também 4 de cada equipe—Equipe D (P13–P16) e Equipe E (P17–P20). E, para o Caso C, contou-se com outros 4 estudantes (dos 6 integrantes da equipe, 66.66%), indicados como P21–P24, Equipe F.

É importante destacar que grande parte dos estudantes estava tendo sua primeira experiência trabalhando com profissionais da indústria apesar de já terem realizado entregas em outros contextos, por exemplo, em bolsa de Iniciação Científica ou em *hackathons* de desenvolvimento de software. Os estudantes que participaram das sessões de *focus group* estavam cursando, quando da realização do estudo, disciplinas entre o segundo e o quarto nível do curso de Bacharelado de algum curso de computação (por exemplo, Engenharia de Software, Sistemas de Informação). Nenhum deles havia aprendido sobre BDD como parte de disciplinas curriculares. Ainda, todos os estudantes estavam também cursando sua graduação de maneira remota, como parte da Pandemia Covid-19.

Para facilitar a compreensão, os resultados são organizados baseado nas perguntas realizadas durante o *focus group*, sendo que a questão 5 (Q5) é a principal contribuição deste artigo. Como previamente mencionado, as demais questões provêm contexto e detalhes sobre a adoção do BDD para um entendimento mais abrangente.

5.1 Descrição do Produto (Q1)

Caso 1. Os participantes das equipes A, B e C relataram que o produto em desenvolvimento tratava-se de um aplicativo de rede social que visa a aproximação dos colaboradores da ORG – “É uma rede social que os colaboradores utilizam para postar, curtir e comentar postagens de outros colaboradores, ler notícias, campanhas” (P1). Outro participante comentou que no aplicativo é possível confirmar a participação em eventos que estão ocorrendo na ORG “Durante a vacinação do Covid-19, os membros do meu time utilizaram o aplicativo para confirmar que iriam na sede da ORG se vacinar, para podermos nos ver pessoalmente mesmo que por uns minutos e com todas as restrições” (P4). Eles ainda comentaram sobre as funcionalidades de caronas e rotas de ônibus – “Quando retornarmos ao trabalho presencial iremos poder utilizar as funcionalidades de oferecer e solicitar caronas e verificar as rotas do ônibus que a empresa disponibiliza” (P7). Os participantes também demonstraram empolgação em utilizar as *features* que eles estavam desenvolvendo –

“Não vejo a hora de usar o app que ajudei a desenvolver. Quero que meus colegas também utilizem e me dêem feedback” (P9).

Caso 2. Os participantes deste caso definiram o produto da seguinte forma: “Nós desenvolvemos um sistema de gestão eletrônica de documentos de recursos humanos.” (P18). P14 complementou dizendo que – “O produto serve para aumentar a produtividade dos funcionários do RH facilitando todo o gerenciamento dos documentos, principalmente na busca”. O participante P16 afirmou – “O software desenvolvido tem a principal função de ajudar as pessoas na gestão de vários tipos de documentos”.

Caso 3. Apesar de serem projetos diferentes, os participantes da Equipe F descreveram o produto em desenvolvimento de forma semelhante aos participantes do Caso 2 – “O software serve para auxiliar pessoas em uma organização a trabalharem”. O participante P21 completa, especificando – “É um portal para os gestores tomarem decisões estratégicas. Eles vão se basear nas informações de outros órgãos e no painel que vai ter na parte inicial do portal”.

Apesar de descrições sucintas, pode-se perceber que todas as equipes perceberam o principal valor agregado esperado do produto em desenvolvimento sob sua responsabilidade. Cabe ressaltar que todas as equipes entregaram o produto no escopo e prazo esperados.

5.2 Adoção do BDD e Ferramentas/Framework (Q2 e Q3)

Caso 1. Os participantes comentaram que o BDD foi adotado no apoio das *user stories* - US, por meio dos critérios de aceitação que são descritos com a linguagem ubíqua Gherkin – “Algumas US já continham os critérios de aceitação de forma detalhada e eram escritas usando as palavras chaves Given, When e Then. Nós recebemos elas assim e precisamos revisar, entender e voltar a (re)discutir com o cliente” (P2). O mesmo participante complementou – “Outras foram definidas por nós assim que passamos a entender a lógica”. Assim, os participantes destacaram que o uso desta linguagem facilitou o entendimento do que deveria ser implementado em cada US – “Os critérios de aceitação descreviam praticamente o passo a passo do que deveríamos implementar e isso nos ajudou a desenvolver as tarefas de cada user story” (P1). As *user stories* eram disponibilizadas para a equipe através do Microsoft Azure DevOps Server², parte da ferramenta Team Foundation System, usada pela empresa – “Os critérios de aceitação são apresentados a equipe durante a *planning* e os POs nos dão acesso ao projeto no Azure para consultar as US e criar as tarefas necessárias para a implementação” (P6). O participante P12 citou que os critérios de aceitação também ajudaram no desenvolvimento de testes funcionais – “Quando terminávamos uma tarefa e íamos fazer os testes, era só olhar os critérios e nos basearmos no comportamento que eles descreviam para que o teste fosse realizado. A gente basicamente colocava o que estava no Given no Arrange, a parte do When no Act e a parte do Then no Assert e era só sucesso. Esses são os comandos do JUnit para testes unitários”.

Caso 2. Neste projeto, o BDD é usado em todo o ciclo de desenvolvimento conforme a Figura 1 – “O BDD foi adotado para guiar o desenvolvimento de algumas funcionalidades e de testes automatizados. Nós tivemos a reunião de *planning* para definir as *user stories* que seriam desenvolvidas na *sprint*. Depois, nós dividíamos as tarefas para iniciar a implementação dos testes e no final fazer a implementação”

(P15). O participante P13 comentou que logo após a *planning*, a equipe se reunia para discutir sobre os testes e posteriormente iniciar o desenvolvimento. Como ferramenta para a disponibilização das *user stories*, tanto as equipes D e E utilizaram o Trello³ – “Nossas US ficavam disponíveis no Trello e as tarefas também eram criadas lá” (P17). – “Na *planning*, as US eram apresentadas à equipe através do Trello e depois criávamos as tarefas.” (P18). Porém os frameworks utilizados foram diferentes, na equipe D foram utilizados o JBehave e JUnit – “Para a implementação dos testes que foram escritos baseados nos critérios de aceitação, utilizamos o JBehave e o JUnit” (P13). E a equipe E utilizou o Cucumber – “Não nos adaptamos ao layout do JBehave, achamos complicado. Então o coordenador do projeto sugeriu utilizar o Cucumber” (P19). Um dos participantes acrescentou que o Cucumber permite a escrita dos cenários e US em português sem a adição de uma biblioteca – “O Cucumber também nos ajudou na escrita das US e dos cenários, pois já escrevamos em português e isso facilitava nossa comunicação com o cliente” (P20).

Caso 3. Neste caso, onde o BDD também é utilizado em todo o ciclo de desenvolvimento, os participantes relataram o uso do Trello para visualizar as *user stories* e seus critérios de aceitação. Dois participantes (P22, P24) mencionaram que esses critérios eram definidos entre o PO e os analistas de negócio por meio de reuniões. Ao serem questionados sobre a utilização de algum framework, o participante P23 citou que eles utilizam o Behave, pois o portal está sendo desenvolvido em Python. P23 também afirmou que a combinação dos dois auxilia no desenvolvimento. O participante complementou mencionando sobre os testes – “Implementar os testes ficou fácil. Usamos o Selenium para automatizar alguns testes, isso ajuda com que nossas tarefas não voltem com frequência”.

5.3 Dificuldades na Adoção do BDD (Q4)

Os participantes das equipes A, B, C (do Caso 1) e D e E (do Caso 2) relataram sobre a dificuldade de escrever as *user stories* e o longo tempo dedicado a esta tarefa – “No início nós demoramos muito na escrita das *user stories*, pois não tínhamos experiência, mas no decorrer da realização das *sprints* nós evoluímos” (P16). P2, da Equipe A, argumentou – “Foi essencial a colaboração com o PO para entender mais sobre as necessidades do cliente. A segurança estabelecidas nas conversas reduziu as dificuldades de entender a técnica”.

Junto a isso, mencionaram o desafio de estabelecer e entender os critérios de aceitação e os seus respectivos cenários. Por exemplo, o participante P7 da equipe B mencionou – “Teve uma *user story* em que os critérios não estavam detalhados suficientemente e ficou confuso para entender. Como fizemos várias perguntas e não estávamos chegando em um entendimento comum, os POs reescreveram estes critérios de forma mais detalhada, que nos ajudou”. Em especial quando os estudantes eram responsáveis por definir estes critérios e cenários – “... tivemos dificuldade no início em entender como definir os critérios de aceitação e escrever os cenários, porém com os treinamentos e orientações do professor-coordenador a equipe evoluiu e conseguiu discutir com o cliente”, falou P13 da Equipe D.

Outra dificuldade foi saber se os critérios e os cenários definidos representavam o comportamento esperado – “Tivemos receio que os critérios e os cenários definidos estivessem mal escritos. Mas quando as *user stories* foram apresentadas e revisadas com o cliente percebemos

²<https://azure.microsoft.com/pt-br/services/devops/server/>

³<https://trello.com/>

que o entendimento estava correto” (Equipe F, P22). Também se teve a expectativa de uso de ferramentas de apoio—*“Tivemos uma dificuldade com a utilização do framework Jbehave, mas a mudança para o Cucumber solucionou o problema de integração com o ambiente que estávamos usando”* (Equipe E, P18).

Os participantes da equipe F relataram a dificuldade de comunicação entre seus integrantes quando da proximidade da entrega—*“Com a entrega se aproximando, a equipe começou a ficar tensa e introspectiva, quase ninguém falava”* (P24). *“Após a primeira entrega descobrimos que era o medo de não termos especificado corretamente, mas deu tudo certo e passamos a nos dar conta que precisávamos conversar sobre medos como esses”* (P21).

5.4 Identificação dos Benefícios e sua Caracterização na adoção do BDD (RQ5)

Em suas respectivas equipes, os participantes P1, P5 e P10 (Caso 1) e P20 (Caso 2) relataram que no decorrer das sprints perceberam uma evolução da equipe e que ficou mais nítido os benefícios do BDD—*“Na terceira sprint foi bem mais fácil entender os critérios de aceitação e saber o que deveríamos fazer pois a equipe já conversava de forma mais objetiva sobre o que cada um precisava fazer e não ficamos muito tempo apenas debatendo”* (P1). Similarmente, P10 relatou—*“No início era tudo complicado de entender, mas conforme as sprints foram passando ficou mais fácil de entender os processos e como o BDD funciona. A partir do relato dos participantes, percebeu-se que os benefícios passam a ser melhor entendidos no decorrer do uso da abordagem. Estes benefícios e sua caracterização são como seguem.*

Comunicação. Esse foi, recorrentemente, em todos os projetos, um dos primeiros benefícios citados. Os participantes disseram que passaram a se comunicar de maneira mais clara e a entender melhor o que deveria ser realizado em cada US—*“Antes de iniciar a implementação nós conversávamos sobre os objetivos e critérios de aceitação da US. Assim, a gente deixava definido o que cada um iria fazer”* (P5). O participante P7 complementou falando a respeito da otimização do tempo na divisão das tarefas—*“Pegávamos o que estava escrito nos critérios e fazíamos as tarefas”* (P7).—*“Lembro que na primeira sprint tivemos dificuldade em fazer as tarefas, mas conforme aprendemos a usar o BDD começamos a nos comunicar melhor e aproveitar os critérios para as tarefas, pois quando uma tarefa ficava pronta significava que um critério de aceitação estava pronto e a gente estava mais próximo de entregar a tarefa”* (P6).—*“Usamos os critérios de aceitação como ponto de partida para conversar e implementar as tarefas mais rapidamente”* (P5).—*“O pessoal conversava mais ricamente, tudo era bem mais objetivo. Tirávamos dúvidas entre nós mesmos e assim não tínhamos grandes impedimentos em nossas tarefas diárias”* (P11). Outro participante citou que melhorou a comunicação com o PO, reduzindo a necessidade de pedidos de esclarecimento—*“A comunicação com o PO ficou objetiva. Quando tivemos dúvidas, procuramos ele e focamos a conversa para o comportamento descrito nos critérios que gerou a dúvida. Isso otimizou o tempo da equipe e também do PO”* (P23). De forma semelhante, o participante P13 comentou que a comunicação foi muito importante, principalmente porque sua equipe também escrevia as *user stories*—*“A comunicação com toda a equipe era bem importante, principalmente quando nós éramos os responsáveis pela escrita. Tínhamos que escrever o comportamento nos critérios de aceitação para depois desenvolver”* (P15).

Entregas mais rápidas. Três participantes de projetos diferentes mencionaram que o melhor entendimento das *user stories* resultaram em entregas com maior agilidade (mais rápidas)—*“Conforme ganhamos experiência, ficou mais fácil entender como iríamos implementar os critérios de aceitação e assim fazíamos entregas mais rápidas e também com maior qualidade”* (P1).—*“Nossas tarefas eram baseadas nos critérios de aceitação, então eram objetivas e a gente assim que acabava elas também finalizava as user stories”* (P14).—*“Eu percebi que entregamos mais rápido quando na terceira sprint entregamos duas user stories dois dias antes do final da sprint e isso não tinha acontecido antes”* (P23). Os demais participantes concordaram e até lembraram de alguns fatos, como o participante P3 que citou que foi possível iniciar o desenvolvimento antecipadamente.

Maior qualidade. O aumento na qualidade foi relatado de diferentes formas. Por exemplo, um participante do Caso 3 mencionou que o todo o escopo passou a ser entregue—*“Eu só percebi essa melhora quando começamos a entregar todas as user stories, lá pela terceira sprint, e elas não precisaram de ajustes”* (P21). Também, que a adoção do BDD auxiliou na redução do débito técnico—*“Eu confesso que fiz uma comparação entre as primeiras entregas e as demais e consegui perceber que nas últimas sprints quase não tivemos débito técnico. As user stories foram entregues como definidas nos critérios de aceitação”* (P10).—*“Ao comparar as nossas entregas da primeira e segunda sprints com a terceira, percebemos que nossos débitos técnicos foram reduzidos, e que entregamos todas as nossas user stories”* (P9).—Um terceiro aspecto foi percebido, a melhora da realização dos testes da aplicação—*“Também melhoramos na implementação dos testes a partir dos critérios presentes nas user stories”* (P6). Durante a discussão sobre esta característica deste benefício, 7 participantes (P2, P4, P14, P16, P18, P20, e P24) comentaram que o mesmo está interligado com as entregas mais rápidas e a redução do retrabalho.

Redução do retrabalho. Em consequência do ganho em qualidade, os participantes relataram que houve a redução do retrabalho, percebido como a redução de débito técnico—*“Nossas entregas estavam com mais qualidade e eram aprovadas sem a necessidade de criar débitos técnicos para fazer ajustes”* (P3).—*“Como desenvolvemos com maior qualidade, o retrabalho foi reduzido, pois o número de débitos técnicos diminuiu”* (P5).—Diretamente relacionado, está o fato da qualidade das entregas ter melhorado—*“Nosso débito técnico diminuiu quando passamos a ter maior qualidade [das entregas]”* (P21). Outro participante citou que o retrabalho foi reduzido somente depois de algumas sprints—*“Não quero desanimar o pessoal que vem por aí, mas é preciso lembrar que na primeira e segunda sprint tivemos muito retrabalho e isso só mudou depois que aprendemos a olhar de maneira mais eficiente para os critérios [de aceitação] e a guiar o desenvolvimento na íntegra por eles”* (P15).

Implementação dos testes. Além dos benefícios já citados, os participantes destacaram que um dos principais benefícios foi a criação do hábito de implementar testes, sendo estes a partir do TDD ou implementados somente após o desenvolvimento como nas equipes A, B e C—*“O direcionamento, principalmente na criação dos testes automatizados, foi essencial”* (P10). Ainda, facilitou o entendimento do escopo do caso de teste—*“Falei sobre isso anteriormente, mas destaco de novo porque para mim foi o principal benefício. Os comportamentos descritos nos critérios de aceitação foram de grande ajuda na implementação dos testes, pois só era transportar o comportamento do cenário para o caso de teste. Ficou fácil entender o que precisava ser*

testado” (P12). Os participantes das equipes D, E e F realizaram todo o ciclo BDD e a implementação dos testes foi realizada antes do desenvolvimento. Estes participantes mencionaram que os comportamentos descritos nas *user stories* facilitaram na implementação do código – “Escrever os testes antes com base nos critérios de aceitação nos ajudou a entender o que realmente precisávamos implementar” (P16). – “A implementação dos testes acabou se tornando rotina, então antes de qualquer coisa verificávamos se existia teste implementado e, se não existisse, fazíamos a implementação para começar o desenvolvimento. Isso nos ajudou na redução do retrabalho e na qualidade do nosso código e do produto de forma geral” (P14). – “Praticamente, pegávamos o que escrevemos no Behave e transcrevíamos para os testes e depois íamos para a implementação” (P22).

Mudanças seguras. Este benefício foi relatado apenas pelas equipes D, E e F. Os participantes citaram que, quando ocorriam mudanças devido a uma nova necessidade do produto, o mapeamento do que precisa ser modificado era mais rápido e dinâmico – “Teve algumas vezes que o PO pediu alterações no sistema e como utilizamos o Cucumber e fizemos os testes antes de implementar, ficou fácil encontrar o que deveríamos fazer para realizar essas mudanças” (P17). Também foi destacada a facilidade de identificar o que precisava ser alterado – “No momento que soubemos das alterações eu meio que entrei em pânico porque não lembrava onde deveríamos mexer, porém o P20 me lembrou do Cucumber e assim que acabou a reunião verificamos lá e fazer as mudanças foi bem tranquilo” (P18). – “O PO solicitou uma mudança e no início achamos que teríamos dificuldades, porém só precisamos mapear no Cucumber e mudar o que era preciso” (P20). Os participantes P22 e P24 fizeram comentários semelhantes aos de P17 e P18. Eles comentaram que apesar da realização de poucas mudanças, quando elas foram necessárias, ocorreu de forma dinâmica, rápida e segura. Os participantes das equipes A, B e C ao serem questionados sobre a ocorrência de mudanças no escopo afirmaram que não houve pedidos de alteração, apenas débitos técnicos criados pelas equipes.

6 DISCUSSÃO

Os resultados do nosso estudo de caso múltiplo, resumidos na Tabela 3, majoritariamente, corroboraram resultados anteriores de estudos científicos em diferentes contextos: melhora na comunicação ([4] [12]), entregas mais rápidas ([4]), maior qualidade ([17]), redução no retrabalho ([17]) e implementação de testes ([17]), apesar deste último aspecto ser apontado como dificuldade no estudo de Nascimento *et al.* [12]. Os resultados também corroboram, em sua maioria, conhecimento disponível na *grey literature*, como o livro de Smart [21] e a revisão de literatura de Couto, Marczak e Gomes [5]. Entretanto, estes estudos não caracterizam os benefícios identificados, primeiro passo no aprofundamento do entendimento de como esses benefícios são conceitualizados para que se possa, no futuro, definir indicadores de qualidade para eles. O foco em equipes novatas na adoção do BDD revelou que estas equipes já conseguem perceber estes benefícios apesar da sua falta de expertise. Isso sugere que é possível de fato se almejar a melhoria do produto e de processos a partir da implementação do BDD.

A caracterização dos benefícios, entretanto, revela que possíveis indicadores a serem derivados dos aspectos levantados precisarão ser melhor explorados para poderem se tornar quantitativos. Por

exemplo, as equipes perceberam que a comunicação ficou mais clara e objetiva. Esta é uma percepção de ordem qualitativa. Outras possuem potencial de discussão da sua implicação, como é o exemplo da percepção da redução do volume de reuniões para esclarecimento de dúvidas. Reduz-se com isso a alocação de tempo para essas reuniões, mas será que a redução em si das reuniões impactou na melhoria da qualidade dos cenários escritos? Existe uma relação de correlação? Estes são próximos passos de investigação na nossa pesquisa rumo à definição dos indicadores.

Também se observou nos resultados uma similaridade de opinião entre os casos e as equipes, caracterizando diferentes formas de adoção do BDD: no apoio ao ciclo completo do desenvolvimento de software ou apenas em uma etapa, como a especificação das *user stories*. Isso corrobora o argumento de Oliveira e Marczak [13] da importância de ter-se cenários BDD bem definidos e de alta qualidade. Os cenários são a base do ciclo BDD.

Mesmo nas equipes onde o PO não se engajava na definição dos critérios de aceitação, identificaram-se resultados similares, sugerindo que profissionais de TI podem passar a melhor entender as necessidades dos usuários ao estruturarem a forma de especificar essas necessidades, com o uso da linguagem Gherkin.

Dado que essas são equipes novatas, pode-se perceber como os integrantes relataram recorrentemente a importância de receberem mentoria e terem sido capacitados. Dessa forma, recomenda-se que iniciativas similares usufruam dessas estratégias e garantam que os aprendizes reflitam sobre sua formação e a adoção de novas práticas, ferramentas, abordagens.

É importante, entretanto, dar atenção ao conceito de testes no contexto do BDD. O conceito de TDD foca em teste unitário, para verificar que o código desenvolvido está correto, bem escrito. Já o conceito de automação de testes do BDD visa facilitar o processo de verificação da implementação de uma funcionalidade (ou necessidade de usuário). Esses conceitos são distintos e notou-se que os estudantes confundiram, vez ou outra, esses conceitos. Como para os Casos 2 e 3 os estudantes haviam recebido especificamente treinamento em TDD e a disciplina de Verificação e Validação de Software é mais adiante na formação do bacharelado de todos os participantes do *focus group*, é inerente e até aceitável essa confusão. Para o Caso 1, o pesquisador discutiu com os estudantes essa distinção e auxiliou os mesmos, durante suas atividades, a buscar o entendimento na prática dado que neste Caso o pesquisador atuou junto dos estudantes como mentor em tempo integral. Fica a recomendação deste item ser monitorado em equipes novatas independente da formação recebida previamente.

7 CONSIDERAÇÕES FINAIS

Esse estudo de caso múltiplo teve como objetivo identificar como são caracterizados os benefícios percebidos na adoção do BDD por equipes novatas no uso dessa abordagem. O estudo com seis equipes de três casos identificou benefícios já citados pela *grey literature*, provendo evidências científicas a esses resultados. O fato destas equipes serem novatas no uso do BDD sugere que mesmo em processo de aprendizado é possível perceber melhorias decorrentes da adoção desta abordagem, sugerindo seu potencial de melhoria do produto de software e de processos adotados pelas equipes de software. Ainda, este estudo de caso revelou aspectos que caracterizam

Tabela 3: Síntese dos benefícios identificados e sua caracterização na adoção do BDD

Benefício	Caracterização
Melhora da comunicação	Conversas mais claras e objetivas Melhor entendimento das <i>user stories</i> Otimização do tempo Redução de reuniões para esclarecimento de dúvidas
Entregas mais rápidas	Ciclos mais rápidos de entrega Iniciar o desenvolvimento antecipadamente
Maior qualidade	Entrega de todas as US propostas para a <i>sprint</i> Redução de débitos técnicos Realização dos testes de funcionalidade
Redução do retrabalho	Redução de débitos técnicos Aumento na qualidade das entregas
Implementação de testes	Criação de teste baseados no critérios de aceitação Criação de rotina de implementar testes (antes do TDD)
Mudanças seguras	Mudanças mais rápidas e dinâmicas Facilidade de identificar o que deveria ser modificado

os benefícios, nunca apresentado por literatura anterior, provendo indícios iniciais de referência para a definição de indicadores de qualidade da adoção do BDD.

Apesar das limitações do estudo, inerentes a um estudo de caso, acredita-se que a escolha de múltiplos casos com diferentes contextos (empresa privada vs órgão governamental, diferentes domínios de negócio, práticas distintas de desenvolvimento de software, etc), favoreça a busca pela generalização de resultados para essa adoção em equipes novatas no uso de BDD. Também, espera-se que esses resultados sigam de inspiração para profissionais e pesquisadores avançarem no entendimento e validação dos benefícios da adoção do BDD e de como mensurá-los pois, como sugerido por Wohlin et al. [25], a academia deve apoiar a indústria de software na validação de métodos, técnicas, processos e ferramentas, usufruindo de seu conhecimento experimental e sistematizado.

AGRADECIMENTOS

Estima-se a colaboração dos participantes, em especial pelo cenário de pandemia dado o Covid-19 vivenciado no mundo. Esta pesquisa é patrocinada pela Dell Brasil usando incentivos da Lei da Informática Brasileira (Lei no. 8.2.48, ano 1991). Sabrina Marczak agradece o CNPq pela Bolsa de Produtividade (Processo 313181/2021-7).

REFERÊNCIAS

- [1] D. North. 2006. [n. d.]. Introducing BDD. (March 2006). Retrieved April, 2020 from <https://dannorth.net/introducing-bdd/>.
- [2] A. Baldissera. 2012. Pesquisa-Ação: Uma Metodologia do “Conhecer” e do “Agir” Coletivo. *Sociedade em Debate* 7 (2012), 5–25.
- [3] L.R. Baskerville. 1999. Investigating Information Systems with action Research. *Communications of the Association for Information Systems* 2, 1 (1999).
- [4] L. P. Binamungu, S. M. Embury, and N. Konstantinou. 2018. Maintaining Behaviour-Driven Development Specifications: Challenges and Opportunities. In *Proceedings of the International Conference on Software Analysis, Evolution and Reengineering, IEEE, Campobasso, Italy*. pp. 175–184.
- [5] T. Couto, S. Marczak, and F. Gomes. [n. d.]. On the Understanding of How to Measure the Benefits of Behavior-Driven Development Adoption: Preliminary Literature Results from a Grey Literature Study. In *Proceedings of the Brazilian Symposium on Software Quality, ACM, São Luís, Brazil, 2020*, pp. 1–7.
- [6] B. B. França, T. Ribeiro, P. S. dos Santos, and G. Travassos. 2015. Using Focus Group in Software Engineering: Lessons Learned on Characterizing Software Technologies in Academia and Industry. In *Proceedings of the Ibero-American Conference on Software Engineering*. Springer, Lima, Peru, pp. 351–364.
- [7] V. Garousi, M. Felderer, and M. V. Mäntylä. 2019. Guidelines for Including Grey Literature and Conducting Multivocal Literature Reviews in Software Engineering. *Information and Software Technology* 106 (2019), 101–121.
- [8] M. Irshad, R. Britto, and K. Petersen. 2021. Adapting Behavior-Driven Development (BDD) for Large-Scale Software Systems. *Journal of Systems and Software* 177 (2021), pp. 110944.
- [9] S. Kamalakar, S. H. Edwards, and T. M. Dao. 2013. Automatically Generating Tests from Natural Language Descriptions of Software Behavior. In *Proceedings of the International Conference on Evaluation of Novel Approaches to Software Engineering*. SciTePress, Angers, France, 238–245.
- [10] B. Malinowski. [n. d.]. *Argonautas do Pacífico Ocidental*. In: Abril Cultural, 1978.
- [11] R. Mugridge. 2003. Test-Driven Development and the Scientific Method. In *Proc. of the Agile Development Conference*. IEEE, Salt Lake City, EUA, pp. 47–52.
- [12] N. Nascimento, A. R. Santos, A. Sales, and R. Chanin. [n. d.]. Behavior-Driven Development: A Case Study on Its Impacts on Agile Development Teams. In *Proceedings of the International Conference on Software Engineering Workshops*. ACM, Seoul, Republic of Korea, 2020, pp. 109–116.
- [13] G. Oliveira and S. Marczak. 2018. On the Understanding of BDD Scenarios’ Quality: Preliminary Practitioners’ Opinions. In *Proceedings of the Requirements Engineering: Foundation for Software Quality (Lecture Notes in Computer Science)*, Vol. 10753. Springer, Utrecht, The Netherlands, pp. 290–296.
- [14] G. Oliveira, S. Marczak, and C. Moralles. 2019. How to Evaluate BDD Scenarios’ Quality?. In *Proceedings of the Brazilian Symposium on Software Engineering*. ACM, Salvador, Brazil, pp. 481–490.
- [15] M. Patton. [n. d.]. *Qualitative Research and Evaluation Methods*. Sage, Thousand Oaks, USA, 2002, pp. 598.
- [16] R. Reilly. 2010. Participatory Case Study. In *Encyclopedia of Case Study Research*, A. Mills and G. Durepos (Eds.). Sage, Thousand Oaks, USA, Chapter 2, pp. 658–661.
- [17] F. Rocha, L. Souza, T. Silva, and G. Rodriguez. 2019. Agile Teaching Practices: Using TDD and BDD in Software Development Teaching. In *Proc. of the Brazilian Symposium on Software Engineering*. ACM, Salvador, Brazil, pp. 279–288.
- [18] P. Runeson and M. Höst. 2009. Guidelines for Conducting and Reporting Case Study Research in Software Engineering. *Empirical Software Eng.* 14, 131 (2009).
- [19] J. Saldaña. 2013. *The Coding Manual for Qualitative Researchers*. Sage, Thousand Oaks, USA, pp. 366.
- [20] A. Silva. 2014. Uma Abordagem Ágil para Transformar Modelos Cognitivos em Modelos Comportamentais e de Domínio. Univ. Nova de Lisboa, Portugal.
- [21] J. Smart. 2014. *BDD in Action: Behavior-Driven Development for the Whole Software Lifecycle*. Manning Publications Shelter Island, pp. 384.
- [22] C. Solis and X. Wang. 2011. A Study of the Characteristics of Behaviour Driven Development. In *Proceedings of the International Conference on Software Engineering and Advanced Applications*. IEEE, Oulu, Finland, pp. 383–387.
- [23] P. Vogt. 1999. *Dictionary of Statistics and Methodology: A Nontechnical Guide for the Social Sciences*. Sage, Thousand Oaks, USA.
- [24] S. Williams and J. Keady. 2021. *Participatory Case Study Work: What it is, How it Functions and Our Adaptations to the Approach*. Routledge, London, UK, pp. 268.
- [25] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, and B. Regnell. 2012. *Experimentation in Software Engineering*. Springer, New York, USA, pp. 236.