

# Usando pesquisas para construir evidências empíricas sobre Dívida Técnica Relacionada ao Teste

## ABSTRATO

**Antecedentes:** Dívida Técnica (DT) refere-se ao problema de tarefas e artefatos pendentes ou incompletos que trazem ganhos de curto prazo para um projeto de software, mas que podem ter que ser pagos com juros posteriormente em seu ciclo de vida. Usamos o termo dívida relacionada a testes para nos referirmos a itens de TD enfrentados por profissionais de teste de software. **Objetivo:** Investigar, do ponto de vista dos praticantes de teste de software, os tipos de TD associados à dívida relacionada ao teste, as causas que a levam e os efeitos de sua existência. **Método:** Realizamos uma replicação de uma pesquisa TD bem conhecida com profissionais de teste de software e combinamos as respostas obtidas com as coletadas no estudo original. No total, o estudo analisa, quantitativa e qualitativamente, 54 respostas de praticantes brasileiros de testagem. **Resultados:** os itens de dívida relacionados ao teste estão associados

com tipos de TD, como documentação, código e dívida de teste em si. A pesquisa identificou 61 causas de DT que estão organizadas em oito categorias. As principais causas são prazo, desatualização/documentação incompleta e planejamento inadequado. A pesquisa identificou 45 efeitos de TD que são agrupados em seis categorias. Os principais efeitos são atraso na entrega, retrabalho e baixa qualidade externa. **Conclusão:** Este artigo organiza as causas e efeitos identificados em diagramas probabilísticos de causa e efeito, que fornecem uma visão ampla do estado da prática em relação às causas e efeitos da dívida relacionada a testes.

## CONCEITOS DE CCS

• Geral e referência ỹ Estudos empíricos • Software e sua engenharia ỹ Manutenção de software

## PALAVRAS-CHAVE

Dívida técnica, dívida relacionada ao teste, teste de software, causas da dívida técnica, efeitos da dívida técnica, diagramas probabilísticos de causa e efeito, InsightTD

## Formato de referência ACM:

Duplo-cego. 2020. Usando Pesquisas para Construir Evidências Empíricas sobre Dívida Técnica Relacionada a Testes. No XXXIV Simpósio Brasileiro de Engenharia de Software (SBES 2020), 19 a 23 de outubro de 2020, Natal, Brasil. ACM, York, NY. <https://doi.org/10.1145/1234567890> EUA, 10 Páginas.

## 1. Introdução

A *dívida técnica (TD)* refere-se ao problema de tarefas e artefatos pendentes ou incompletos que trazem ganho de curto prazo para um

projeto de software, mas pode ter que ser pago com juros mais tarde em seu ciclo de vida [1, 2]. Em projetos de software, o teste é uma atividade importante para garantir a qualidade dos sistemas de software [3]. No entanto, esta atividade também pode sofrer com a presença de itens de dívida.

Falta de testes, testes adiados e falta de planejamento de casos de teste são exemplos de situações que apontam para a existência de dívida de teste [4]. Analisando essas situações, podemos perceber que elas estão relacionadas apenas a artefatos e tarefas produzidas ou executadas por praticantes de teste de software. No entanto, outros artefatos e tarefas podem influenciar na ocorrência de TD nas atividades de teste. Por exemplo, especificações de requisitos desatualizadas (uma dívida de requisitos) podem produzir casos de teste incorretos. Nesse cenário, usamos o termo *dívida relacionada a testes (TRD)* para nos referirmos a problemas enfrentados por profissionais de teste de software devido à presença de itens TD.

Uma maneira de lidar com itens de TRD é entender seus *tipos de TD*, as *causas* que levam as equipes de software a incorrer neles e os *efeitos* de sua presença. Ter informações sobre os efeitos do TRD auxilia na priorização dos itens do TRD a serem compensados, apoiando uma análise de impacto mais precisa e a identificação de ações corretivas para minimizar possíveis consequências negativas desses itens para o projeto. Por outro lado, conhecer as causas do TRD pode auxiliar as equipes de desenvolvimento na definição de ações preventivas.

Vários pesquisadores têm investigado as causas e efeitos da DT [5-16]. Por exemplo, Ernst et al. [5] avaliaram uma lista predefinida de causas de TD pesquisando engenheiros e arquitetos de software e Martini e Bosh [7, 8] investigaram os efeitos da dívida de arquitetura realizando um estudo de caso em cinco e quatro grandes empresas, respectivamente. Rios et al. [11, 16] iniciou um esforço em larga escala chamado de *Projeto InsightTD*, uma família globalmente distribuída de pesquisas industriais, para investigar as causas e efeitos da TD em projetos de software. Nenhum dos trabalhos citados foca sua discussão especificamente nas atividades de teste de software.

O objetivo do nosso trabalho é investigar, do ponto de vista dos praticantes de teste de software, os tipos de TD associados à *dívida relacionada ao teste*, as causas que a levam e os efeitos de sua existência. Para isso, utilizamos dados de TRD coletados pelo Projeto *InsightTD* [16], doravante denominado estudo original, e também dados coletados por nossa própria replicação do *InsightTD*, focado em praticantes de teste de software. No total, 54 profissionais de teste de software do Brasil (12 do estudo original e 42 de nossa replicação) responderam à pesquisa.

Nossos resultados constroem um corpo de conhecimento sobre TRD, que está associado a doze tipos de dívida, principalmente *dívida de teste*, *dívida de documentação* e *dívida de código*. De um total de 61 TRD

causas, as mais citadas são: *prazo, documentação desatualizada/incompleta e planejamento inadequado*. De um total de 45 Efeitos TRD, os mais citados são: *atraso na entrega, retrabalho e baixa qualidade externa*. As causas e efeitos são agrupados em categorias e relacionados aos tipos de TD. Assim como Rios e cols. [16], nosso trabalho também organiza seu corpo de conhecimento em um conjunto de diagramas probabilísticos de causa-efeito, mas focados no ponto de vista de praticantes de teste de software.

O restante deste trabalho está organizado da seguinte forma. A Seção 2 apresenta uma breve introdução ao projeto *InsightTD*. A seção 3 apresenta o método de pesquisa. A seção 4 apresenta os resultados obtidos. A seção 5 apresenta os diagramas de causa e efeito probabilísticos para TRD. A Seção 6 compara nossos achados com o trabalho relacionado. A Seção 7 resume nossas descobertas. A seção 8 discute as implicações do estudo para pesquisadores e profissionais. A seção 9 apresenta ameaças à validade do estudo. Por fim, a Seção 10 apresenta nossas considerações finais.

## 2 O Projeto InsightTD

O projeto *InsightTD* foi concebido como uma família de pesquisas distribuída globalmente para identificar o estado da prática e as tendências do setor em TD. Seu desenho permite a replicação da pesquisa em diferentes países. Até o momento, pesquisadores do Brasil, Chile, Colômbia, Costa Rica, Finlândia, Índia, Itália, Noruega, Arábia Saudita, Sérvia e EUA aderiram ao projeto.

Vários trabalhos já foram publicados a partir das análises dos dados do *InsightTD*. O estudo original [16], baseado nos primeiros resultados publicados em Rios et al. [11], apresentam o projeto de pesquisa e o conjunto geral de causas e efeitos da DT identificados nos dados originalmente coletados no Brasil. Rios et al. [15] ir mais longe nesse conjunto de dados e analisar dados relacionados a projetos de software ágil no Brasil. Para organizar o corpus de conhecimento das causas e efeitos da DT, Rios et al. [12] propuseram o uso de diagramas probabilísticos de causa-efeito de TD entre empresas.

Outras iniciativas do projeto *InsightTD* focaram a gestão de TD. Freire et al. [17] apresentam uma lista de ações preventivas de DT e impedimentos que dificultam a aplicação dessas ações. Freire et al. [18] apresentam uma lista de práticas de pagamento de TD e discute os motivos da não aplicação dessas práticas. Pérez et al. [19] centram-se na relação entre as práticas de pagamento e os efeitos da DT.

Replicações da pesquisa foram realizadas por Pacheco et al. [20], Pérez et al. [13], e Ramac et al. [14], na Costa Rica, Chile e Sérvia, respectivamente. O trabalho de Pacheco et al. [20] apresenta a caracterização dos participantes da pesquisa e como eles reagem à presença de itens TD. Os demais trabalhos discutem as causas e efeitos da DT encontrados e os comparam com os identificados pelo estudo original [16].

Em nosso artigo, focamos em analisar os dados do *InsightTD* considerando apenas os praticantes de teste de software para investigar as causas e os efeitos do TD percebidos por eles. Conduzimos nossa própria replicação brasileira, aplicada apenas a testes de software

profissionais, e complementou com os dados de TRD coletados no estudo original [16]. Nossa replicação é diferente das demais [13, 14, 20], pois não ocorreu em outro país e é focada apenas em TRD.

## 3 Método de Pesquisa

Nesta seção, apresentamos as questões de pesquisa definidas neste trabalho e discutimos sua metodologia.

### 3.1 Questões de Pesquisa

Definimos a seguinte questão principal de pesquisa (QR) **“Como os profissionais de software percebem a ocorrência de itens de dívida relacionados a testes em atividades de teste de software?”**. Esta questão está dividida nas seguintes subquestões:

**RQ1:** *Quais são os tipos de TD presentes nas atividades de teste?*

Esta questão visa investigar os tipos de TD relacionados às atividades de teste. Vamos mais longe no conjunto de tipos de TD e derivamos a seguinte subquestão:

**RQ1.1:** *Existe alguma correlação entre os tipos de DT identificados? Por meio dessa pergunta, buscamos identificar correlações entre os tipos de DT. Essa correlação pode indicar que quando um tipo de TD é encontrado em atividades de teste, seus tipos de TD correlacionados também podem ser encontrados.*

**RQ2:** *Quais são as principais causas de dívidas relacionadas a testes citadas por profissionais de teste de software em seus projetos?* O objetivo desta pergunta é identificar as causas que levam as equipes de desenvolvimento a incorrer em TRD. Derivamos a seguinte subquestão para verificar a relação entre as causas e o tipo de TD:

**RQ2.1:** *Quais são as principais causas de TRD citadas por profissionais de teste de software por tipo de TD?* Esta questão visa reconhecer as causas relacionadas a cada tipo de DT identificado no RQ1.

**RQ3:** *Quais são os principais efeitos do TRD citados pelos profissionais de teste de software em seus projetos?* Esta questão pretende identificar os efeitos do TRD sentidos pelas equipes de desenvolvimento nos seus projetos. Também temos uma subquestão para o RQ3:

**RQ3.1:** *Quais são os principais efeitos do TRD citados pelos profissionais de teste de software por tipo de TD?* Esta questão visa reconhecer os efeitos relacionados a cada tipo de DT identificado no RQ1.

### 3.2 Coleta de Dados

Este trabalho utiliza a pesquisa definida no projeto *InsightTD*, mas duas estratégias de coleta foram realizadas para derivar nosso conjunto de dados. O primeiro empresta dados do estudo original [16], enquanto este último coleta novos dados usando nossa própria replicação.

Embora a pesquisa seja composta por 28 perguntas, focamos neste trabalho em um subconjunto dessas perguntas, apresentadas na Tabela 1 juntamente com o número identificador da pergunta (**Não**), descrição (**Pergunta (Q)**) e digite (**Tipo**). Q1 a Q8 são *questões de caracterização*. Q13 e Q15 pedem aos participantes que descrevam um exemplo de um item TD que ocorreu em seu projeto de software e indiquem o nível de representatividade deste exemplo, respectivamente. Por último, o 16º e o 19º trimestre focam-se no

discussão das causas do TD, enquanto Q20 e Q21 conduzem a discussão sobre os efeitos do TD.

A estratégia de coleta de dados foi o envio de convites por e-mail, ou use canais de convite, como LinkedIn, para alcançar grupos afiliados ao setor, listas de discussão e parceiros do setor.

Todas as respostas coletadas foram validadas seguindo dois critérios de aceitação: (i) o participante era um praticante de teste de software, e (ii) o exemplo dado no Q13, na verdade, descrevia um item TD. Esses critérios garantem que os dados coletados estejam alinhados ao nosso objetivo de pesquisa.

3.3 Análise de Dados

Como o questionário é composto por questões fechadas e abertas (ver Tabela 1), diferentes estratégias de análise de dados foram empregadas no estudo. Para as questões fechadas, utilizamos estatística descritiva e calculamos a proporção de participantes que escolheram cada opção. Essas estratégias foram utilizadas nas questões de caracterização.

Aplicamos a análise de dados qualitativos às perguntas abertas [21, 22]. Nas respostas dadas à Q13, usamos a lista de indicadores de TD fornecida em [23] para **identificar o tipo de TD**. A lista de tipos de TD reconhecidos e sua frequência suportam a resposta para RQ1. Considere a seguinte resposta para Q13 como um exemplo de identificação do tipo TD: *“Falha ao usar testes de unidade e má escolha de arquitetura de software”*. Este exemplo revela problemas em teste e arquitetura e, como tal, sinaliza dívida de teste e dívida de arquitetura.

Para responder a RQ1.1, quando uma resposta à Q13 mencionava mais de um tipo de TD, utilizamos uma matriz de correlação e interpretamos seus valores levando em consideração a seguinte categorização de Hopkins [24]: (i) quase perfeito (0,9 a 1,0, ou - 0,9 a -1,0), muito grande (0,7 a 0,9 ou -0,7 a -0,9), grande (0,5 a 0,7 ou -0,5 a -0,7), moderado (0,3 a 0,5 ou -0,3 a -0,5), pequeno

(0,1 a 0,3 ou -0,1 a -0,3) e insubstancial (menor que 0,1 e maior que -0,1). Também aplicamos o teste de correlação de Kendall em cima disso. Este teste é indicado para distribuições não normais e amostras pequenas. Nossos dados se enquadram nessas indicações porque: (i) confirmamos que nossos dados não seguem uma distribuição normal por meio do teste de normalidade de Shapiro-Wilk e (ii) identificamos mais de um tipo de DT em apenas 12 respostas.

Com base nas respostas dadas a Q16-Q21, aplicamos **a codificação manual** seguindo uma abordagem lógica indutiva [21]. Inicialmente, o conjunto de todas as respostas foi codificado para quatro subconjuntos. Dois subconjuntos representam as causas do DT, e foram formados pelas respostas dadas à Q16-Q19 e pelas respostas dadas à Q16-Q18. O primeiro subconjunto suporta responder RQ2 enquanto o segundo suporta responder RQ2.1. Os outros dois subconjuntos são compostos de efeitos TD. A primeira é formada pelas respostas dadas a Q20 e Q21, corroborando a resposta RQ3 e a outra pelas respostas dadas à Q21, corroborando a resposta à RQ3.1. Finalmente, códigos com mesmo significado foram unificados. Este processo foi realizado de forma iterativa até atingir um ponto de saturação, onde não foram identificados novos códigos.

A codificação manual exigiu algum esforço para identificar os códigos e realizar sua unificação. Por exemplo, os participantes citaram as seguintes causas de TD: *“falta de teste unitário”, “falta de teste funcional pelo fornecedor”* e *“não execução de testes”*. Essas causas de TD estão relacionadas a (i) falta de testes unitários, (ii) falta de testes funcionais e (iii) testes não executados, respectivamente. Ao final, foram codificados como *teste não realizado*.

Como muitos dos códigos de um mesmo subconjunto estavam relacionados entre si, seguimos um **processo de agrupamento** para organizá-los em categorias [21, 22]. Quando um pesquisador analisava os códigos e identificava uma relação entre eles, ele os agrupava em uma categoria. Para nomear os grupos, utilizamos as categorias

Tabela 1: Subconjunto das perguntas da pesquisa *InsightTD* relacionadas às causas e efeitos da TD (adaptado de [16])

Não.	Pergunta (Q)	Modelo
Q1	Qual o tamanho da sua empresa?	Fechadas
Q2	Em qual país você está trabalhando atualmente?	Fechadas
Q3	Qual é o tamanho do sistema que está sendo desenvolvido nesse projeto? (LOC)	Fechadas
Q4	Qual é o número total de pessoas deste projeto?	Fechadas
Q5	Qual é a idade deste sistema até agora ou quando seu envolvimento terminou?	Fechadas
Q6	A qual função do projeto você está atribuído neste projeto?	Fechadas
Q7	Como você avalia sua experiência nesta função?	Fechadas
Q8	Qual das opções a seguir descreve mais de perto o modelo de processo de desenvolvimento que você segue neste projeto?	Fechadas
Q13	Dê um exemplo de TD que teve um impacto significativo no projeto que você escolheu para nos contar: Q15 Sobre este exemplo, quão representativo ele é?	Abrir Fechadas
Q16	Qual foi a causa imediata ou precipitante do exemplo de DT que você acabou de descrever?	Abrir
Q17	Que outra causa ou fator contribuiu para a causa imediata que você descreveu acima?	Abrir
Q18	Que outros motivos ou razões ou causas contribuíram direta ou indiretamente para a ocorrência do exemplo TD?	Abrir
Q19	Considerando todos os casos de TD que você encontrou em diferentes projetos, e as causas desses casos de TD, quais causas você diria são os mais propensos a levar a TD (ordenados pela probabilidade de causar TD)? Por favor, liste até 5 causas.	Abrir
Q20	Considerando o item TD, você descreveu na questão 13, quais foram os impactos sentidos no projeto?	Abrir
Q21	Considerando todos os casos de TD que você encontrou em diferentes projetos e os efeitos desse TD que você pessoalmente experimentados, quais 5 efeitos você classificaria como os efeitos de maior impacto (ordenados por seu nível de impacto).	Abrir

Tabela 2: Funções dos participantes do teste de software

Função	#FR	%FR	Nível de Experiência			
			Iniciante (9%)* 0	Competente (52%) 3	Proficiente (32%) 3	Especialista (7%)
Analista de teste	6	11%				0
Gerente de teste/testador	46	85%	3	25	14	4
Testador	2	4%	2	0	0	0

\* O número entre parênteses representa a porcentagem do total de participantes com esse nível de experiência.

proposto pelo estudo original [16]. Por exemplo, usamos a categoria *Problemas de qualidade externa* para agrupar efeitos de TD como *baixa qualidade externa*, *projeto não atende às expectativas do cliente* e *baixo desempenho*.

Todos os processos descritos nesta seção foram realizados por primeiro e segundo autores, separadamente. As inconsistências foram resolvidas em uma reunião de consenso. Além disso, na replicação, os pesquisadores que realizaram a codificação manual utilizaram as listas de causa e efeito publicadas pelo estudo original [16] para assegurar a consistência com aquele estudo.

4 Resultados

No total, obtivemos 54 respostas válidas coletadas de profissionais de teste de software no Brasil: 12 do estudo original [16] e 42 de nossa replicação.

4.1 Caracterização do Teste de Software Praticantes

A Tabela 2 apresenta os papéis dos participantes, relatando o nome do papel (**papel**), a frequência de ocorrência de cada papel (**#FR**) e o percentual de FR em relação ao total de todos os papéis citados (**%FR**). A coluna **nível de experiência** mostra o número de participantes por seu nível de experiência em sua função entre as seguintes opções: Iniciante (conhecimento prático dos principais aspectos da prática), Competente (bom conhecimento prático e de fundo da área de prática), Proficiente (profundidade de compreensão da disciplina e área de prática) e Expert (conhecimento autoritário da disciplina e profundo entendimento tácito em toda a área de prática). Para cada nível de experiência, apresentamos, entre parênteses, sua porcentagem em relação a todo o conjunto de dados.

A maioria dos entrevistados se identificou como Test Manager/Testador (85%), seguido por Test Analyst (11%) e Testador (4%). A maioria dos respondentes se identificou como competente (52%), seguido de proficiente (32%), iniciante (9%) e especialista (7%), indicando que, em geral, o questionário foi respondido por profissionais de teste de software com experiência em sua área de atuação. funções.

Organizações de diferentes portes compuseram o conjunto de dados. A maioria são empresas de médio porte (59%, organizações com 51 a 1.000 funcionários), seguidas de grandes (37%, mais de 1.000 funcionários) e pequenas (4%, até 50 funcionários). Em relação ao processo utilizado pelos entrevistados, 55% era ágil, 30% era híbrido (uma combinação de métodos ágeis e tradicionais

práticas), e 15% era tradicional. Quanto ao tamanho da equipe, a maioria dos entrevistados trabalhava em equipes com mais de 30 pessoas (33%), seguidas de 10-20 pessoas (30%), 5-9 pessoas (26%), 21-30 pessoas (7 %) e menos de 5 pessoas (4%). A idade do sistema mais mencionada nos exemplos de TD foi superior a dez anos (26%), seguida de 1 e 2 anos (23%), mas também tivemos vários sistemas com 2 a 5 e 5 a 10 anos (ambos com 20%) e menores de 1 ano (11%).

O tamanho de sistema mais comum foi 10-100 KLOC (35%), entre 100-1000 KLOC (24%), 10+ MLOC (20%), menos de 10 KLOC (13%) e 1-10 MLOC (8% ).

Em relação às respostas dadas à Q13, a maioria representa (Q15) um tipo de coisa que acontece com muita frequência no projeto (48%), seguido de um tipo de coisa que acontece de vez em quando (37%) e uma instância única (15%).

Em resumo, o conjunto de dados apresenta diferentes perfis de praticantes brasileiros de teste de software, considerando sua função, nível de experiência, organização e tamanho da equipe, modelo de processo e idade e tamanho do sistema.

4.2 Quais são os tipos de TD presentes nas atividades de teste? (RQ1)

A Tabela 3 mostra os tipos de TD identificados em exemplos dados por profissionais de teste de software no Q13, relatando o número de ocorrências (**#TDT**) de um tipo de TD e a porcentagem de **#TDT** (**%TDT**) em relação ao total de respostas válidas. Encontramos 12 tipos de TD relacionados a atividades de teste. De toda a lista de tipos de DT listados por Rios et al. [4], apenas dívida de infraestrutura, dívida de usabilidade e dívida de versionamento não foram identificadas.

Tabela 3: Tipos de TD identificados nas atividades de teste

NÃO	Tipo de dívida técnica	#TDT	%TDT
1º	Teste de dívida	20	37%
2º	Dívida de documentação	15	28%
3º	Código da dívida	7	13%
4º	Dívida com defeito	5	9%
5º	Dívida do projeto	4	7%
6º	Dívida obrigatória	4	7%
7º	Dívida de teste de automação	3	6%
8º	Processar dívida	3	6%
9º	Dívida de arquitetura	2	4%
10º	Dívida das pessoas	2	4%
11º	Construir dívida	1	2%
12º	Dívida de serviço	1	2%

Rubrica:  
NÃO - Ordem numérica  
#TDT - Número de ocorrências no exemplo TD (Q13)  
%TDT - Porcentagem de #TDT sobre o total de respostas válidas

Entre os tipos de dívida identificados, a dívida de teste (37%) é a mais comum, seguida da documentação (28%), código (13%) e dívida de defeito (9%). Por outro lado, arquitetura (4%), pessoas (4%), construção (2%) e dívida de serviço (2%) são os menos comuns nos projetos.

4.2.1 Existe alguma correlação entre os tipos de DT identificados?  
(RQ1.1).

Encontramos mais de um tipo de TD em 12 exemplos (Q13), as linhas destacadas em cinza na Tabela 3 indicam esses tipos. Calculamos o coeficiente  $\gamma$  de Kendall para cada combinação de tipos. A Figura 1 mostra esses coeficientes organizados em uma matriz de correlação e o efeito de cada coeficiente seguindo a categorização de Hopkins [24]. Pode-se notar que a dívida de código e de teste de automação (0,95) tem uma correlação quase perfeita entre eles, indicando que se a dívida de código é encontrada nas atividades de teste, a dívida de teste de automação também pode ser encontrada nessas atividades e vice-versa. O efeito também é muito grande ou grande, para dívida de documentação e requisito (0,73), dívida de defeito e pessoas (0,59) e dívida de teste e processo (0,56). Isso significa que, se um desses tipos de TD ocorrer, o outro também poderá ocorrer em atividades relacionadas a testes.

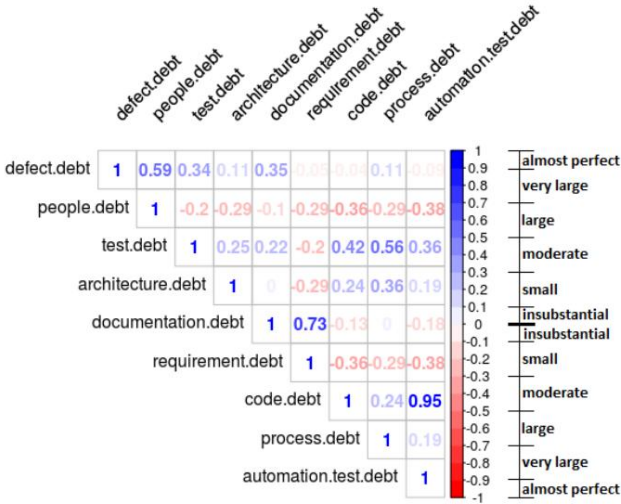


Figura 1: Matriz de Correlação entre os tipos de TD

Para entender melhor essas correlações e seus efeitos, verificamos ainda se essas correlações são significativas entre cada par de tipo de TD. Como os dados não seguem uma distribuição normal, foi realizado o teste de correlação de Kendall com nível de confiança de 95% para cada par de tipo TD. Apenas os pares código e dívida de teste de automação ( $p$  valor = 0,0055), e documentação e dívida de requisitos ( $p$  valor = 0,0248) tiveram correlações significativas, significando que o tipo de TD em cada par pode ocorrer em conjunto nas atividades de teste. O resultado completo dos testes de normalidade e correlação está disponível em <https://bit.ly/35wHBC9>.

4.3 Quais são as principais causas de TRD citadas pelos profissionais de teste? (RQ2)

Do 16º ao 19º trimestre, identificamos 61 causas associadas a dívidas relacionadas a testes. A Tabela 4 apresentou a lista das 5 principais causas, relatando o nome da causa, o total de citações para a causa (#C) e o percentual de #C em relação ao total de causas citadas (%C). No total, as 5 principais causas correspondem a ~28% da frequência geral de citações. A lista completa de causas de dívidas relacionadas ao teste está disponível em <https://bit.ly/35wHBC9>.

Tabela 4: As 5 Principais Causas de Dívida Relacionadas a Testes Citadas

Causa da dívida relacionada ao teste	#C	%C
Data limite	25	8,3%
Documentação desatualizada/incompleta	17	5,7%
Planejamento inadequado	15	5,0%
Falta de conhecimento técnico	14	4,7%
Teste não realizado	12	4,0%

Rubrica:  
#C - Número de citações para a causa  
%C - Porcentagem de #C em relação ao total de causas citadas

Seguindo o processo de agrupamento (ver Subseção 3.3), organizou as causas de dívida relacionadas ao teste em oito categorias:

- **Questões de desenvolvimento:** engloba problemas que ocorrem durante o desenvolvimento do projeto. *Código ruim, falta de qualidade externa e não adoção de boas práticas* são exemplos das quinze causas que compõem esta categoria.
- **Fatores externos:** inclui causas relacionadas a problemas externos à equipe de desenvolvimento e à organização. *Pressão, cliente não ouve a equipe do projeto e componente descontinuado* são exemplos das três causas que compõem esta categoria.
- **Infraestrutura:** agrupa causas associadas a problemas de infraestrutura. Apenas as causas *necessárias* *infraestrutura indisponível, problemas com ambiente de teste e escolha de tecnologia/ferramenta/plataforma inadequada* são agrupados nesta categoria.
- **Falta de conhecimento:** tem causas relacionadas ao desconhecimento da equipe para desenvolver o projeto. *Falta de conhecimento técnico, falta de experiência e falta de conhecimento de tecnologia* são exemplos das seis causas que compõem esta categoria.
- **Metodologia:** engloba causas relacionadas a processos e metodologias utilizadas no desenvolvimento do projeto. *Documentação desatualizada/incompleta, teste não realizado e falta de um processo bem definido* são exemplos das doze causas que compõem esta categoria.
- **Organizacional:** inclui causas relacionadas ao nível de organização da equipe. *Alta rotatividade da equipe, falta de equipe específica e cultura organizacional* são exemplos das quatro causas dessa categoria.
- **Pessoas:** grupos de causas relacionadas a membros de equipes de desenvolvimento de software. *Falta de comunicação da equipe, falta de*

Tabela 5: Causas de dívida relacionadas a testes mais citadas por tipo de TD

Causa da dívida relacionada ao teste	Tipo de TD											
	TD	DoD	CD	DeD	DfD	RD	ATD	PrD	AD	PeD	BD	SD
Data limite	7	2	2	1	1	1	1	1	0	0	0	0
Documentação desatualizada/incompleta	4	3	1	0	0				0	0	0	0
Planejamento inadequado	5	2	2	3	0	1	1	0	0	0	0	0
Falta de conhecimento técnico	1	2	0		1	0	0	0	0	1	0	0
Teste não realizado	2	0	0	1	0	0	0	0	0	0	0	0
Rubrica:	TD - Teste de dívida			DeD – Dívida de projeto			ATD - Dívida de teste de automação			PeD - Dívida de pessoas		
	DoD - Dívida de documentação			DfD - Dívida com defeito			PrD - Processar dívida			BD -Construir dívida		
	CD - Código de débito			RD - Dívida Requerida			AD - Dívida de arquitetura			SD - Dívida de serviço		

comprometimento e sobrecarga da equipe são exemplos das sete causas que compõem esta categoria.

- **Planejamento e gestão:** tem causas associadas ao planejamento e gestão de projetos. *Prazo, planejamento inadequado e estimativa de tempo imprecisa* são exemplos das onze causas que compõem essa categoria.

4.3.1 Quais são as principais causas de dívidas relacionadas a testes citadas por profissionais de teste de software por tipos de TD? (RQ2.1).

A partir das causas de TRD identificadas no Q16-Q18 e do tipo de TD identificado no Q13, derivamos a relação entre as causas e os tipos de TD. A Tabela 5 apresenta o número de ocorrências de cada uma das 5 principais causas para cada um dos tipos de DT identificados. Podemos observar que, exceto arquitetura, pessoas, construção e dívida de serviço, o *prazo* de causa motiva o surgimento dos outros tipos de TD. A dívida de teste é motivada por todas as 5 principais causas citadas, enquanto nenhuma delas motiva o surgimento da dívida de arquitetura, construção e serviço. A lista completa de causas de dívidas relacionadas ao teste para cada tipo de TD está disponível em <https://bit.ly/35wHBC9>.

4.4 Quais são os principais efeitos do TRD citados pelos profissionais de teste? (RQ3)

De Q20 e Q21, identificamos 45 efeitos associados ao TRD. A Tabela 6 apresenta a lista dos 5 principais desses efeitos, informando o nome do efeito, o número total de citações para o efeito (**#E**) e a porcentagem de #E em relação ao total de efeitos citados (**%E**). No total, os 5 principais efeitos correspondem a ~42% da frequência geral de citações. A lista completa dos efeitos da dívida relacionados ao teste está disponível em

<https://bit.ly/35wHBC9>.

Organizamos os efeitos da dívida relacionados ao teste em seis categorias:

- **Questões de desenvolvimento:** engloba efeitos relacionados às atividades de desenvolvimento do projeto. *Testes inadequados, documentação inadequada/inexistente/desatualizada e não implementação de funcionalidade* são exemplos dos dez efeitos que compõem esta categoria.
- **Problemas de qualidade externa:** inclui efeitos relacionados à qualidade ao longo do desenvolvimento de software. Apenas os efeitos *de baixa qualidade externa, o projeto não se encaixa no cliente*

expectativas e *baixo desempenho* são agrupados nesta categoria.

- **Problemas internos de qualidade:** agrupa os efeitos associados às atividades de manutenção de software. *Baixa manutenibilidade, necessidade de refatoração e código ruim* são exemplos dos oito efeitos que compõem esta categoria.
- **Organizacional:** tem efeitos relacionados à organização.

Apenas os seguintes efeitos compõem esta categoria: *prejuízo financeiro e questões legais por não cumprimento de contratos*.

- **Pessoas:** efeitos de grupos que dizem respeito a questões relacionadas aos membros da equipe de desenvolvimento. *A insatisfação dos stakeholders, a desmotivação da equipe e o estresse com os stakeholders* são exemplos dos treze efeitos que compõem essa categoria.
- **Planejamento e gestão:** engloba os efeitos relacionados ao planejamento e gestão de projetos. *Atraso na entrega, retrabalho e maior esforço* são exemplos dos nove efeitos que compõem essa categoria.

Tabela 6: Os 5 principais efeitos de dívida relacionados a testes citados

Efeito da dívida relacionada ao teste	#E	%E
Atraso de entrega	24	11,3%
Retrabalho	23	10,8%
Baixa qualidade externa	17	8,0%
Insatisfação das partes envolvidas	14	6,6%
Perda financeira	11	5,2%
Rubrica:		
#E - Número de citações para o efeito		
%E - Porcentagem de #E em relação ao total de efeitos citados		

4.4.1 Quais são os principais efeitos da dívida relacionada a testes citados por profissionais de teste de software por tipos de TD? (RQ3.1).

A partir dos efeitos identificados nos tipos Q20 e TD reconhecidos no Q13, identificamos a relação entre os efeitos TRD e os tipos TD. A Tabela 7 apresenta o número de ocorrências dos tipos de DT e os efeitos mais citados. Podemos notar que o *atraso na entrega* do efeito é sentido devido a itens de teste, documentação, código, defeito, projeto, requisito e dívida de processo. A dívida de teste produz todos os 5 principais efeitos citados, enquanto nenhum deles é produzido por dívida de construção e serviço. A lista completa de efeitos de dívida relacionados ao teste por tipo de TD está disponível em <https://bit.ly/35wHBC9>.

Tabela 7: Efeitos da dívida relacionados ao teste mais citados por tipo de TD

Efeito da dívida relacionada ao teste		Tipo de TD											
		TD DoD CD DeD				DfD				PrD			
Atraso de entrega		6	6		1	2	1					0	0
Retrabalho		10	3	1	0	0	0	0	1	1	0	0	0
Baixa qualidade externa		1	1	2	0	1	0	0	0	0	0	0	0
Insatisfação das partes envolvidas		3	2	0	0	2	0	0	0	0	0	1	0
Perda financeira		2	0	0	0	2	0	0	0	0	0	0	0
Rubrica:	TD - Teste de dívida	DeD - Dívida de projeto				ATD - Dívida de teste de automação				PeD - Dívida de pessoas			
	DoD - Dívida de documentação	DfD - Dívida com defeito				PrD - Processar dívida				BD - Construir dívida			
	CD - Código de débito	RD - Dívida Requerida				AD - Dívida de arquitetura				SD - Dívida de serviço			

## 5 Diagramas Probabilísticos de Causa-Efeito para Dívida Relacionada ao Teste

Organizamos as informações sobre causas e efeitos da dívida relacionada ao teste em diagramas probabilísticos de causa-efeito. Este tipo de diagrama foi proposto e avaliado em [12], demonstrando que são úteis para subsidiar análises de causas e efeitos da DT.

As Figuras 2 e 3 mostram os diagramas definidos para as causas e efeitos do TRD, respectivamente. Cada diagrama é composto pelos seguintes elementos: (i) um retângulo preto com bordas arredondadas denominado "Dívida Técnica Relacionada ao Teste" posicionado à esquerda (efeito) ou à direita (causa) e uma linha principal chegando (causa) ou saindo (efeito) este retângulo; (ii) um conjunto de retângulos conectados à linha principal por linhas secundárias. Cada retângulo representa uma categoria, contendo seu nome, sua probabilidade, e é pintado em tom cinza. Quanto maior a probabilidade, mais escuro o cinza; e (iii) em cada linha secundária, um conjunto de ramificações representa causas ou efeitos, contendo o nome e a probabilidade dos mesmos.

Para chegar aos valores percentuais, usamos a lista final de

causas (Q16-Q19) e efeitos (Q20 e Q21). Para isso, somamos o número de vezes que cada causa/efeito foi citado.

Em seguida, calculamos a porcentagem para cada causa/efeito dividindo o número total de citações de cada causa/efeito pelo número total de citações de causas/efeitos. Seguimos o mesmo processo para definir o percentual em cada categoria.

Na Figura 2, podemos observar que as principais causas que levam ao endividamento relacionado ao teste estão concentradas na categoria *planejamento e gestão* (26% das vezes). Dentro dessa categoria, a causa mais citada é o *prazo* (8,33%), seguido de *planejamento inadequado* (5%) e *estimativa de tempo imprecisa* (3,33%). Na Figura 3, podemos observar que os principais efeitos que impactam nas atividades de teste estão concentrados na categoria *planejamento e gerenciamento* (~33% das vezes).

Dentro dessa categoria, o efeito mais citado é o *atraso na entrega* (11,32%), seguido pelo *retrabalho* (10,85%) e *aumento do esforço* (4,25%).

Também definimos um conjunto de diagramas TRD probabilísticos de causa-efeito para cada tipo de TD. Os diagramas estão disponíveis em <https://bit.ly/35wHBC9>.

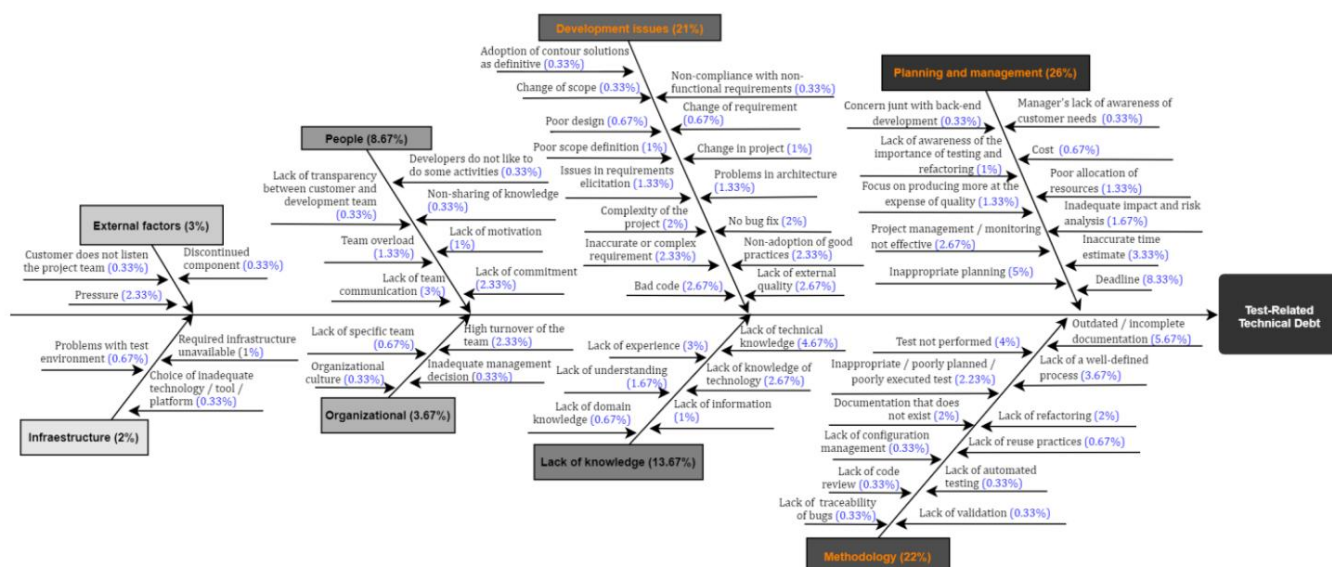


Figura 2: Diagrama de causa e efeito probabilístico para causas de dívida relacionadas a testes



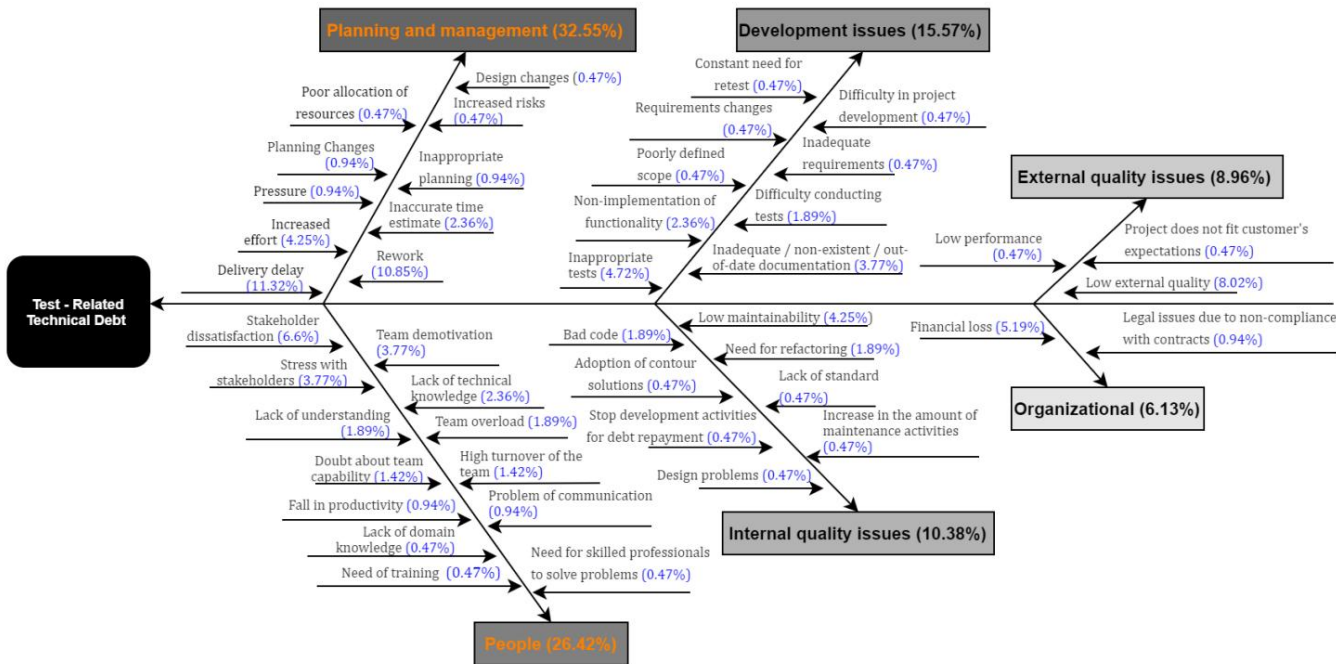


Figura 3: Diagrama de causa e efeito probabilístico para efeitos de dívida relacionados ao teste

6 Comparação com Trabalho Relacionado

Vários trabalhos de pesquisa investigaram as causas e efeitos da DT [5-10]. No entanto, as evidências relatadas ainda são limitadas em relação ao tamanho da amostra, ao modelo de processo adotado, ao número de organizações diferentes ou ao uso de uma lista pré-definida de causas de DT, direcionando as respostas dos participantes. O projeto *InsightTD* foi iniciado para mitigar essas limitações [11-16]. Em particular, o estudo original, baseado em estudos preliminares [11,12], realizou uma pesquisa no Brasil coletando 107 respostas de praticantes de software considerando seu papel, nível de experiência, organização e tamanho da equipe, modelo de processo e idade e tamanho do sistema. Nosso estudo vai na mesma direção, mas focamos especificamente em dívida relacionada a testes, buscando evidências empíricas sobre TRD na indústria brasileira.

A Figura 4 mostra dois diagramas de Venn comparando as causas (Figura 4-A) e os efeitos (Figura 4-B) listados no estudo original e em nosso estudo. Identificamos apenas duas novas causas e nenhum efeito novo em nossa replicação. Isso era esperado, pois o estudo original se concentrou em DT em geral, enquanto nosso estudo se concentrou em TRD. No entanto, nosso estudo adiciona contexto a essas informações: agora sabemos quais causas e efeitos são relacionadas a atividades de teste de software. Portanto, os resultados obtidos em nosso trabalho refletem mais de perto os problemas e suas causas enfrentados pelos praticantes de teste de software.

7 Resumo das Descobertas

As principais conclusões deste trabalho podem ser resumidas nos seguintes pontos:

- Doze tipos de TD estão associados a dívidas relacionadas a testes. Entre eles, dívida de teste, dívida de documentação e dívida de código são os principais tipos encontrados nas atividades de teste;
- Nas atividades de teste, dívida de código e dívida de teste de automação podem ser encontrados juntos. O mesmo pode ocorrer com a dívida de documentação e a dívida de exigência;
- Prazo, documentação desatualizada/incompleta e planejamento inadequado são as causas de TRD mais citadas;
- Atraso na entrega, retrabalho e baixa qualidade externa são os efeitos da dívida relacionados aos testes mais citados;
- As principais causas que levam ao TRD e os principais efeitos que impactam nas atividades de teste estão concentrados no planejamento e gestão da categoria;
- A categoria questões de desenvolvimento concentra o maior número de causas de TRD, enquanto a categoria pessoas o maior número de efeitos de TRD. Quanto aos efeitos, isso é um indicativo de que o TRD tem um papel importante nos aspectos humanos do desenvolvimento de software;
- Um conjunto de diagramas probabilísticos de causa e efeito foi definido organizando as causas e efeitos da dívida relacionada ao teste. Além disso, definimos diagramas para cada tipo de TD associado à dívida relacionada ao teste.

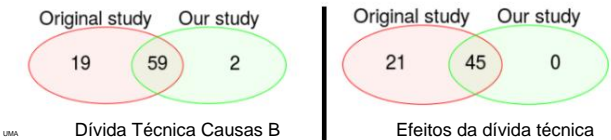


Figura 4: Comparação com o Estudo Original e Nosso estudo considerando as Causas do TD (A) e Efeitos do TD (B)



Usando pesquisas para construir evidências empíricas sobre testes relacionados  
Dívida Técnica

SBES 2020, 19 a 23 de outubro de 2020, Natal, Brasil

## 8 Implicações para Profissionais e Pesquisadores

As equipes de desenvolvimento podem usar nossa lista de causas e efeitos de dívidas relacionadas a testes para entender como os itens TD surgem em seus projetos e seus possíveis efeitos. Os diagramas de causa-efeito representam o estado da prática em termos de causas e efeitos de TRD. Esses diagramas são construídos especificamente para os praticantes de teste e podem apoiá-los melhor do que os diagramas de propósito geral propostos em [16].

Os praticantes de teste de software podem usar os diagramas em reuniões sobre TD para reconhecer causas e efeitos de dívida relacionados a testes em geral ou por tipo de TD, para que possam propor ações para conter itens de TD. Por exemplo, se uma equipe de software identifica que *as mudanças de requisito* (uma causa associada à dívida de requisito) são responsáveis por incorrer em itens de TD nas atividades de teste, a equipe pode analisar (i) a lista de efeitos de dívida relacionados ao teste ou (ii) a lista de efeitos relacionados à dívida de requisito para entender o impacto dessa causa no projeto.

Para os pesquisadores, nossos resultados apoiam pesquisas futuras, fornecendo insights sobre as perspectivas dos profissionais de teste de software sobre as causas e efeitos da TD. As listas de causas e efeitos permitem que os pesquisadores guiem suas pesquisas de uma forma orientada para o problema.

## 9 Ameaças à Validade

Utilizamos as categorias definidas por Wohlin et al. [25] identificar as ameaças à validade do nosso trabalho, procurando eliminá-las ou mitigá-las.

A principal ameaça à **validade de construto** surge da seleção dos participantes. Como pretendemos investigar os tipos de TD, causas e efeitos associados à dívida relacionada a testes, apenas os profissionais de teste de software devem responder ao questionário. Para reduzir essa ameaça, usamos respostas dadas por profissionais de teste de software coletadas do estudo original e da replicação.

Os processos utilizados para identificação de DT, codificação manual e agrupamento representam ameaças à **validade da conclusão** devido à sua subjetividade e possível inconsistência. Para mitigar essa ameaça, cada processo foi realizado por pelo menos dois pesquisadores separadamente, e reuniões de consenso entre eles foram realizadas para dirimir divergências. Outra ameaça de conclusão está relacionada ao tamanho da amostra utilizada no teste de correlação.

Apesar de termos manuseado uma pequena amostra, consideramos o característica da amostra na escolha do teste estatístico apropriado e seguiu cuidadosamente suas premissas.

Em relação à **validade interna**, o questionário *InsightTD* representa uma ameaça à validade interna, pois as questões podem não ser claras e objetivas para todos os participantes. Para reduzir essa ameaça, o grupo original realizou três validações internas, uma validação externa e um estudo piloto antes da primeira execução do questionário [16].

Finalmente, uma importante ameaça de **validade externa** está relacionada ao tamanho e perfil da população. Nossa amostra parece fornecer uma boa diversidade de praticantes de teste de software da indústria brasileira, mas o número final de respostas não é suficiente para garantir a generalização dos resultados obtidos. Outras replicações no Brasil são necessárias para aumentar a validade externa das respostas.

## 10 Observações Finais

Este trabalho aborda o ponto de vista dos praticantes sobre a dívida relacionada ao teste, ou seja, uma dívida encontrada nas atividades de teste. Ele fornece insights sobre dívida relacionada a testes, incluindo tipos de TD associados a ela e suas principais causas e efeitos. O corpo de conhecimento produzido foi organizado em diagramas probabilísticos de causa-efeito, facilitando a compreensão do estado da prática em relação às causas e efeitos do TRD.

Os próximos passos desta pesquisa incluem: (i) investigar se as causas e efeitos do TRD são influenciados pela experiência do participante, modelo de processo adotado e tamanho da organização/projeto, (ii) realizar novas replicações no Brasil para aumentar sua validade externa, (iii) compreender a situação da dívida relacionada a testes em outros países e (iv) avaliar os diagramas de causa e efeito com relação à sua eficácia no apoio à gestão da dívida relacionada a testes.

## REFERÊNCIAS

- [1] Clemente Izurieta, Antonio Vetrò, Nico Zazworka, Yuanfang Cai, Carolyn Seaman e Forrest Shull, 2012. Organizando o cenário da dívida técnica. In *Proceedings of 3rd MTD*, Zurique, p. 23-26.
- [2] Rodrigo O. Spínola, Nico Zazworka, Antonio Vetrò, Forrest Shull e Carolyn Seaman, 2019, Compreendendo abordagens de identificação de dívida técnica automatizada e baseada em humanos: um estudo de duas fases. *Revista da Sociedade Brasileira*, 25(5).
- [3] Javaria Imtiaz, Salman Sherin, Muhammad Uzair Khan e Muhammad ZohaibIqbal, 2019. Uma revisão sistemática da literatura sobre técnicas de prevenção e reparo de quebra de teste. *Tecnologia da Informação e Software*, 113 (9), setembro, p. 1-19.
- [4] Nicoll Rios, Manoel Mendonça Neto e Rodrigo O. Spínola, 2018. Um estudo terciário sobre dívida técnica: tipos, estratégias de gestão, tendências de pesquisa e informações de base para praticantes. *Tecnologia da Informação e Software*, 102(6), p. 117-145.
- [5] Neil A. Ernst, Stephany Bellomo, Ipek Ozkaya, Robert L. Nord e Ian Gorton. 2015. Medir? Gerenciá-lo? Ignore isto? praticantes de software e dívida técnica. Em *Anais da 10ª ESEC/FSE*. pág. 50-60.
- [6] Antonio Martini, Jan Bosch e Michel Chaudron. 2014. Dívida técnica da arquitetura: entendendo as causas e um modelo qualitativo. In *Proceedings of the 40th EUROMICRO SEAA*, Verona, Itália, p. 85-92.
- [7] Antonio Martini e Jan Bosch. 2015. O perigo da dívida técnica arquitetônica: dívida contagiosa e círculos viciosos. In *Proceedings of the 12th Working IEEE/IFIP Conference on Software Architecture*, Montreal, p. 1-10.
- [8] Antonio Martini e Jan Bosch. 2015. Priorizando a dívida técnica da arquitetura: necessidades de informação de arquitetos e proprietários de produtos. In *Actas da 41ª Euromicro SEAA*, Funchal, p. 422-429.
- [9] Jesse Yli-Huoma, Andrey Maglyas e Kari Smolander. 2014. As fontes e abordagens para a gestão da dívida técnica: um estudo de caso de duas linhas de produtos em uma empresa finlandesa de software de médio porte. Em *Anais do XV PROFES*, p. 93-107.
- [10] Jesse Yli-Huoma, Andrey Maglyas e Kari Smolander. 2015. Os benefícios e consequências das soluções alternativas em projetos de desenvolvimento de software. Em *Anais do 6º ICSBO*, p. 1-16.
- [11] Nicoll Rios, Rodrigo Oliveira Spínola, Manoel Mendonça e Carolyn Seaman. 2018. As causas e efeitos mais comuns da dívida técnica: primeiro

SBES 2020, 19 a 23 de outubro de 2020, Natal, Brasil

Duplo-cego et al.

- resultados de uma família global de pesquisas industriais. Em *Atas do 12º ESEM*, pág. 1-10.
- [12] Nicolli Rios, Rodrigo O. Spínola, Manoel Mendonça e Carolyn Seaman. 2019. Apoiar a análise de causas e efeitos de dívida técnica com diagramas de causa e efeito probabilísticos entre empresas. Em *Atas da 2ª TechDebt*, p. 3-12.
- [13] Boris Pérez, Juan Pablo Brito, Hernán Astudillo, Darío Correal, Nicolli Rios, Rodrigo O. Spínola, Manoel Mendonça e Carolyn Seaman. 2019. familiaridade, causas e reações de profissionais de software à presença de dívida técnica: um estudo replicado na indústria de software chilena. Em *Anais do 38º SCCC*, p. 1-7.
- [14] Robert Ramaj, Vladimir Mandij, Nebojša Taušan, Nicolli Rios, Manoel Mendonça, Carolyn Seaman e Rodrigo O. Spínola. 2020. Causas e efeitos comuns da dívida técnica na TI sérvia: replicação da Pesquisa InsightTD. Submetido à *46ª SEAA* – disponibilizado pelos autores.
- [15] Nicolli Rios, Manoel Mendonça, Carolyn Seaman e Rodrigo O. Spínola. 2019. Causas e efeitos da presença de dívida técnica em projetos de software ágeis. Em *Anais do AMCIS*.
- [16] Nicolli Rios, Rodrigo O. Spínola, Manoel Mendonça e Carolyn Seaman, 2020. O ponto de vista dos praticantes sobre o conceito de dívida técnica e suas causas e consequências: um projeto para uma família global de pesquisas industriais e seus primeiros Resultados do Brasil. *Engenharia de software empírica Diário*. Aceito para publicação – disponibilizado pelos autores.
- [17] Sávio Freire, Nicolli Rios, Manoel Mendonça, Davide Falessi, Carolyn Seaman, Clemente Izurieta e Rodrigo O. Spínola. 2020. Ações e impedimentos para a prevenção da dívida técnica: resultados de uma família global de pesquisas industriais. In *Proceedings of the 35th ACM SAC*, pp. 1548-1555.
- [18] Sávio Freire, Nicolli Rios, Bóris Pérez, Darío Correal, Manoel Mendonça, Clemente Izurieta, Carolyn Seaman e Rodrigo O. Spínola. 2020. Pesquisando profissionais de software sobre práticas técnicas de pagamento de dívidas e razões para não pagar itens de dívida. In *Anais da 24ª EASE*.
- [19] Bóris Pérez, Camilo Castellanos, Darío Correal, Nicolli Rios, Sávio Freire, Rodrigo Spínola e Carolyn Seaman. 2020. Quais são as práticas utilizadas pelos profissionais de software no pagamento de dívidas técnicas? Resultados de uma família internacional de pesquisas. Em *Anais da 3ª TechDebt Conf*.
- [20] Alexia Pacheco, Gabriela Marín-Raventós e Gustavo López. 2019. Dívida técnica na Costa Rica: uma replicação da pesquisa InsightTD. Em *Anais do XX PROFES*, p. 236-243.
- [21] Carolyn Seaman. 1999. Métodos qualitativos em estudos empíricos de engenharia de software. *IEEE Transactions on Software Engineering*, 25(4):557-572.
- [22] Anselm Strauss e Juliet M. Corbin. 1998. Noções básicas de pesquisa qualitativa: técnicas e procedimentos para o desenvolvimento da teoria fundamentada. Sage Publicações.
- [23] Nicolli SR Alves, Thiago S. Mendes, Manoel G. de Mendonça, Rodrigo O. Spínola, Forrest Shull e Carolyn Seaman, 2016. Identificação e gestão da dívida técnica: um estudo de mapeamento sistemático. *Tecnologia da Informação e Software*, vol. 70, fevereiro, pág. 100-121.
- [24] Will G. Hopkins, "A new view of statistics," Sport Science, 2003. <http://www.sportsci.org/resource/stats/>.
- [25] Claes Wohlin, Per Runeson, Martin Höst, Magnus C. Ohlsson, Björn Regnell e Anders Wesslén. 2012. Experimentação em engenharia de software: uma introdução. Springer.