

# Evaluation of reproducibility and accuracy of the business process point analysis technique

Natália Pereira de Oliveira, Marcelo Fantinato  
*School of Arts, Sciences and Humanities  
University of São Paulo  
São Paulo, Brazil  
Email: n.oliveira, m.fantinato@usp.br*

Lucinéia Heloisa Thom  
*Institute of Informatics  
Federal University of Rio Grande do Sul  
Porto Alegre, Brasil  
Email: lucineia@inf.ufrgs.br*

**Resumo**—Techniques of functional size measurement are easily found in the literature, however, in the evaluation process of these techniques is not always approached which makes its validity questionable. The evaluation of the Business Process Point Analysis (BPPA) technique is the object of study of this article that aims to consistently evaluate its reproducibility and accuracy, identifying its limitations. BPPA was proposed so that project managers can systematically estimate the functional size of a business process automation project. Thus, this article presents the execution of a quasi-experiment realized with 58 graduate and postgraduate students, who measured the functional size of three business process models. The results of this experiment present the low reproducibility and accuracy of the technique as well as its limitations.

**Index Terms**—business process model, reproducibility, accuracy, functional size, complexity

## 1. Introdução

Análise de Pontos de Processo de Negócio (BPPA) é uma técnica proposta para adaptar a Análise de Pontos de Função (APF), da engenharia de software, para a área de Gestão de Processos de Negócio (BPM) [1], [2]. Enquanto APF é aplicada aos requisitos de software, BPPA é aplicada aos modelos de processo especificados em *Business Process Management Notation* (BPMN) [2]. BPPA foi proposta para que gerentes de projeto possam usar técnicas para obter informações relacionadas sobre o tamanho funcional de projetos da área de BPM, sendo capazes de estimar sistematicamente o tamanho funcional de um processo de negócio [1]. Esse procedimento de medição é necessário pois pode auxiliar gerentes de projeto a identificar e gerenciar esforços, custos e cronogramas de processos de negócio [3].

Embora a técnica tenha sido proposta, o processo de definição de medidas definido por Muketha em [4] não foi realizado por completo, pois, não foi submetida a uma avaliação consistente. O processo de avaliação a que foi submetida contempla apenas fatores relacionados à viabilidade, à facilidade e aos benefícios da aplicação [1].

Dessa forma, o objetivo deste artigo é apresentar os resultados de um experimento, visando identificar a capa-

cidade de reprodutibilidade e acurácia da técnica BPPA e assim poder identificar suas limitações. Neste artigo são usadas as definições do Vocabulário Internacional de Metrologia (VIM) para reprodutibilidade e acurácia [5]

- **Reprodutibilidade**, acordo entre os resultados de diferentes sujeitos usando o mesmo a técnica BPPA.
- **Acurácia**, é o acordo entre a proximidade dos resultados dos sujeitos e o valor real.

Para isso, um experimento foi conduzido com 58 sujeitos, obtendo 137 respostas. No experimento, três modelos de processo de negócio foram usados em que cada sujeito teve que aplicar BPPA para os modelos de processo. Para avaliação dos resultados, estatística descritiva foi usada.

O restante deste artigo apresenta: fundamentação teórica, trabalhos relacionados, método de pesquisa, protocolo desenvolvido para a realização do experimento, bem como a classificação dessa pesquisa e os resultados alcançados, ameaças à validade e por fim, a conclusão da pesquisa.

## 2. Fundamentação teórica

Nesta seção são apresentados os conceitos: gestão de processos de negócio, medidas de software e a técnica análise de pontos de processos de negócio.

### 2.1. Gestão de processos de negócio

A gestão de processos de negócio tem por objetivo aproveitar oportunidades de melhoria, assegurando melhores resultados e organizando a realização de trabalho nas organizações. BPM se preocupa com a gestão de cadeias de eventos, por meio de atividades e decisões denominando processos de negócio, com o objetivo de agregar valor às organizações e aos seus clientes [6].

Processos de negócios são um conjunto de atividades realizadas em coordenação em um ambiente organizacional e técnico a fim de atingir uma meta de negócios [7], e podem ser vistos como ativos de organizações, capazes de revelar seu nível de maturidade [8]. Com o uso de BPM na organizações, é possível atingir com mais facilidade o

alinhamento estratégico entre áreas de negócio e a área de tecnologia da informação, gerenciando e aprimorando as soluções tecnológicas a partir de processos que incorporam valor significativo para as organizações.

## 2.2. Medição de software

O termo medida é definido como um instrumento usado para associar um valor de um elemento mensurável usando uma escala para o mapeamento de uma medida [9]. Deste modo, para que um processo de medição seja estabelecido é necessário que as entidades do mundo real a serem medidas sejam identificadas (elemento mensurável) e que as relações numéricas correspondentes sejam identificadas, permitindo que relações empíricas sejam estabelecidas para os elementos mensuráveis de um determinado sistema numérico [10].

As medidas visam avaliar coleções de dados sobre atividades de determinado projeto e recursos. Medidas na indústria de tecnologia da informação são usadas para quantificar recursos e/ou processos de desenvolvimento de software. Os padrões de medição têm sido muito usados nas indústrias de software como uma maneira de avaliar a eficácia e eficiência de atividades do projeto e devem ser usadas com o objetivo de auxiliar na estimativa, no andamento, no desempenho e na quantificação de atributos do projeto [11].

A estimativa de processos de negócios é definida como uma tarefa composta por atribuir, de maneira empírica e objetiva, números às propriedades dos processos de negócios de modo a descrevê-los [12]. Relacionado a processos de negócio, atributos de medição incluem complexidade, custo, capacidade de manutenção e confiabilidade [12]. Neste artigo, a complexidade do processo é referida por tamanho, pois trata-se do valor numérico que representa sua complexidade. Assumimos para este artigo que medição funcional de processos de negócio refere-se a capacidade de mensurar a partir de pontos de processos de negócio o seu tamanho funcional.

A estimativa da complexidade de processos de negócio foi definida da seguinte forma: Seja P um conjunto de processos. Seja N um conjunto numérico. Uma medida m é definida para ser um mapeamento,  $m: P \rightarrow N$ , garantindo que todo processo tenha um valor de estimativa e que cada processo tenha apenas uma medida [12].

## 2.3. Visão Geral da técnica BPPA

A técnica análise de pontos de processo de negócio foi proposta com o objetivo de adaptar a técnica análise de pontos de função, da engenharia de software, a fim ser aplicada em gestão de processos de negócio. BPPA é voltada para aplicação em modelos de processo de negócios modelados BPMN, enquanto o APF da engenharia de software é voltado à requisitos de software. A intenção da proposta de BPPA não busca melhorar a técnica APF, e sim permitir sua aplicação em outro domínio [1].

BPPA foi proposta para que gerentes de projeto de BPM sejam capazes de estimar sistematicamente o tamanho funcional de um processo de negócio a ser usado em um

projeto de automação de processos de negócio. Trata-se da primeira técnica proposta para esse fim conforme avaliação realizada na literatura pelos autores dessa técnica [2].

Buscando a definição sólida de uma técnica que oferecesse esse resultado, BPPA foi definida com base em APF, porém, tendo especificamente como objeto o tamanho funcional de um determinado software.

Enquanto APF é baseada em pontos de função, BPPA é baseada na medida conceitualmente similar, os pontos de processos de negócio (Business Process Points – BPPs). Como resultado da aplicação da técnica BPPA, obtém-se o tamanho funcional de um processo de negócio expresso em BPPs. O cálculo de BPPs é feito com base em um modelo de processo de negócio expresso via BPMN [2]. Embora a técnica BPPA tenha sido baseada em APF, ela não depende de APF para seu uso, a relação entre essas técnicas é apenas a adaptação, como um modelo ou boa prática para BPM. Desse modo, BPPA é vista como uma proposta de uso por gerentes de projetos acostumados com o uso de APF ao trabalharem com BPM [1].

Dessa forma, para conduzir uma contagem de pontos de processos e negócio para projetos de BPM devem ser executadas as mesmas etapas que uma contagem de pontos de função aplicada ao contexto de APF, porém, a documentação do projeto BPM é basicamente composta por modelos de processo, e o BPPA foca apenas nos modelos de processo especificados na notação BPMN versão 2.0.

Os autores desenvolveram um fluxograma, apresentado na Figura 1 que apresenta o fluxo das etapas de BPPA, conforme o diagrama original de FPA.

Assim como no processo de contagem de tamanho funcional de APF a contagem dos BPPs é realizada em dois momentos: na estimativa das funções de dados e na estimativa das funções de transações. Em ambas as fases é necessário analisar o modelo de processos de negócio modelado em BPMN e com base nas estruturas de dados presentes no modelo a contagem é realizada e a complexidade pode ser medida. Os valores padrões usados em APF para esse fim foram mantidos para a contagem para BPM.

Nesse caso, como base para o cálculo três componentes são considerados: (1) arquivos lógicos internos ou arquivos de interface externa, (2) entrada externa e (3) saída externa ou consulta externa. Por meio da complexidade desses componentes, é realizada uma ponderação e a soma dos resultados [2]. A definição completa do cálculo, assim como a variação em relação à técnica FPA são descritas com maiores detalhes em [1], [2], [13].

Duas avaliações empíricas foram realizadas a fim de determinar os benefícios e as desvantagens da BPPA. A primeira é baseada nas observações dos autores ao aplicar a técnica proposta a um cenário de uso específico, e a segunda, se baseia na realização de dois experimentos realizados, um com estudantes e outro com profissionais da indústria [13].

O experimento teve por objetivo avaliar a técnica BPPA, comparando-a com o APF, visando identificar se BPPA é semelhante à técnica APF em termos de facilidade de aplicação, de viabilidade e de potenciais benefícios, uma vez que ambas compartilham os mesmos conceitos básicos

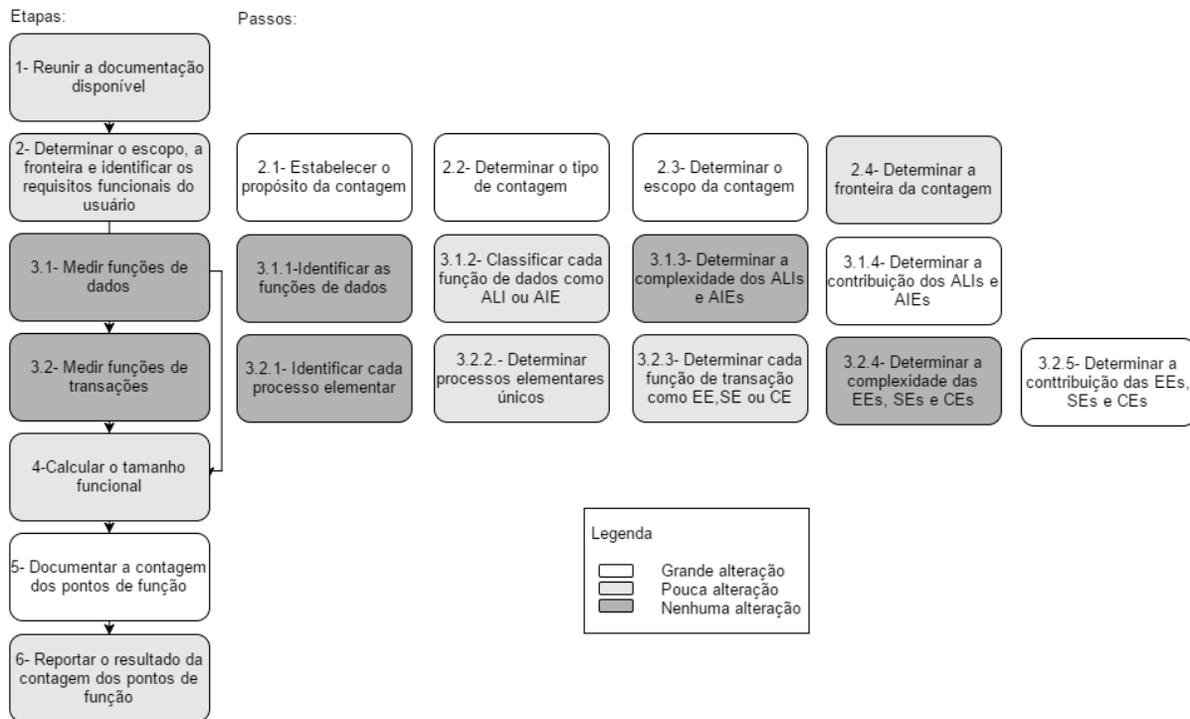


Figura 1. Fluxo de etapas da técnica BPPA [13].

[1]. O experimento teve como hipótese a ideia de que BPPA é próximo a FPA em termos de facilidade de aplicação, viabilidade e benefícios, não necessitando ser melhor, apenas similar [1].

Para a avaliação, os autores usaram o modelo “*The Travel Booking*” [14]. Trata-se de um projeto de desenvolvimento de BPM composto por apenas um PM-BPMN. Com o uso desse modelo, na primeira avaliação, os autores estimaram os BPPs do projeto e obtiveram como resultado, em termos gerais, que a técnica BPPA pode ser aplicada como projetada e proposta pelos autores [1].

Esse estudo foi realizado com 28 estudantes de graduação com antecedentes em BPM e engenharia de software, divididos em duplas a fim de simular ambientes empresariais. Os sujeitos foram submetidos a um treinamento prévio relacionado às técnicas, e aplicaram cada uma das técnicas a três cenários específicos. Após isso, responderam um questionário sobre suas experiências com as técnicas, e sobre suas percepções e comparações entre BPPA e FPA [1]. O questionário continha perguntas com possibilidade de resposta de 0 a 10 em relação à percepção, o valor 5 significava igualdade, abaixo a isso que BPPA era pior que FPA e acima disso que BPPA era melhor que FPA [1].

Estatística descritiva não paramétrica e teste de hipóteses foram realizados. A estatística não paramétrica  $\chi^2$  foi aplicada para identificar se houve diferenças significativas entre os grupos separados em sujeitos experientes, médios e inexperientes. Após a avaliação, o teste de hipótese *Student* foi realizado com um nível de significância de 5%, visando mitigar as ameaças de validade identificadas e vieses [1].

A segunda parte do experimento, com profissionais da indústria, foi realizada de forma independente e complementar em relação ao primeiro estudo, entretanto, visando avaliar os mesmos pontos do experimento feito unos de graduação. Este segundo experimento não foi realizado em ambiente controlado, com treze especialistas certificados em pontos de função com contato diário com BPM e engenharia de software, selecionados por conveniência [1]. Os resultados proporcionaram evidências positivas dos aspectos observados, mostrando que o BPPA está em conformidade com o APF, confirmando a hipótese de pesquisa.

Os resultados proporcionaram evidências positivas dos aspectos observados, mostrando que o BPPA está em conformidade com o FPA, confirmando a hipótese de pesquisa. Os autores sugerem como trabalhos futuros a realização de estudos adicionais: (1) incluindo especialistas imparciais do domínio do BPM, bem como mais especialistas do campo de métricas de software; (2) executar uma experiência em uma configuração do mundo real; (3) aumentar a garantia de validade definindo um conjunto de métricas objetivas a serem aplicadas por meio de um protocolo claro; e (4) conduzir uma abordagem comparativa entre BPPA e outras técnicas relacionadas à medida de BPM [1].

### 3. Trabalhos Relacionados

Abordagens são encontradas na literatura visando realizar a medição funcional de esforços, complexidade, requisitos, entre outros. Nenhum estudo foi encontrado visando realizar a medição do tamanho funcional de processos de

negócio e sua validação. Esta seção apresenta três trabalhos relacionados às técnicas de medição funcional e suas respectivas avaliações.

A técnica de medição *Package Points based ERP Size Estimation* (PPSE) foi desenvolvida a fim de estimar pontos de tarefa, calculados a partir da somatória do número de tarefas e de processos necessários para implementar um requisito, levando em conta sua complexidade. Cada uma das tarefas é avaliada por meio do processo denominado valor de complexidade, realizado com o uso de APF. Por fim, a somatória dos pontos de pacote indica o tamanho do pacote ERP [15]. Catorze projetos ERP desenvolvidos pela empresa Axys foram usados [15], em que os sujeitos calcularam os pontos de pacote para cada um dos projetos, obtendo informações sobre os valores de complexidade e os pontos de tarefa. Foi realizada ainda a estimativa de esforços, usando o *COConstructive COst MOdel* (COCOMO), um modelo de estimativa do tempo de desenvolvimento de um software e a medição de pontos de função com APF. Por fim, a avaliação da precisão da estimativa da técnica PPSE foi realizada, com o uso do método estatístico de correlação [15]. Por fim, os autores mostram que é possível dimensionar pacotes ERP personalizados com base em requisitos de customização (pacotes) em vez de confiar apenas em sua funcionalidade. Os autores afirmam que esses pontos permitiram a PPSE produzir melhores resultados para medição de tamanho de software e estimativa de esforço quando comparado a outros métodos convencionais na prática para projetos de ERP.

Outra abordagem relacionada é a *Object Oriented Hypermedia Function Point* (OO-HFP), que apresenta um procedimento de medição funcional para projetos web usando o método Hipermídia Orientada a Objetos (OO-H). Para tal, o primeiro passo é estimar os números de tipos de elementos de dados, de registro e o número de funções transacionais de registro. Em seguida, as regras do APF devem ser aplicadas para classificar a complexidade da função, em que um valor de pontos de função é atribuído para cada função dependendo de sua complexidade [16]. A avaliação desse método foi realizada por meio da comparação da precisão de predição que a estimativa fornece com a de outro conjunto de medidas de tamanho validados. A pesquisa foi realizada por meio de um *survey*, usando 133 respostas a um formulário web. Uma segunda pesquisa envolvendo 33 empresas virtuais validaram as medidas empiricamente como bons preditores de esforço, e validaram usando o método estatístico Wilcoxon Signed Rankings, com  $\alpha = 0,05$ . A precisão raciocínio baseado em razão não foi significativamente melhor que as previsões com base na mediana ou na média do conjunto de dados [16].

Por fim, o método *OO-Method Function Points* (OomFP) foi desenvolvido para medir o tamanho funcional de sistemas orientados a objetos baseado na medição de esquemas conceituais. Trata-se de um método que mede a visão externa lógica do software sob a perspectiva do usuário, avaliando o número de funcionalidades a serem entregues. Esse método permite quantificar o tamanho funcional dos sistemas orientados a objetos em uma fase inicial do ciclo de vida do desenvolvimento. Esse estudo é conduzido

em comparação da OomFP com APF no que diz respeito a uma série de variáveis baseadas no desempenho e na percepção, incluindo a eficiência (esforço necessário para aplicar os métodos), reprodutibilidade, acurácia, facilidade de uso percebida, utilidade percebida e intenção de uso [17]. A validação da técnica OomFP consistiu em três fases [17]. A primeira fase foi a validação da concepção do método, para o qual foi aplicado um modelo teórico para avaliação de métodos de um projeto de sistemas de informação. Na segunda fase a validação da aplicação do método, ocorreu por meio da aplicação de um experimento laboratorial em um grupo de estudantes de ciência da computação treinados em OO-Method, APF e OomFP. Esta avaliação objetivou analisar “como” (eficácia) e “se” (eficiência) a medição de tamanho funcional de um sistema se torna melhor se o método OomFP for usado em vez de APF, comparando os resultados. A terceira e última fase, ocorreu com a validação do uso dos resultados da medição, que não abrangeu o escopo do trabalho. Esse é o estudo mais relacionado à proposta deste artigo, de modo que a validação realizada objetivou medir a reprodutibilidade e acurácia, que são justamente o foco de avaliação neste estudo. Entretanto, o método empregado para calcular esses fatores não foram compartilhados neste artigo.

#### 4. Método de Pesquisa

O procedimento técnico adotado para fins deste artigo, deriva de um experimento, tratando-se especificamente de um quase-experimento<sup>1</sup>. O experimento executado foi baseado nos princípios de experimentos da engenharia de software definidos por Wohlin [18].

O quase-experimento foi selecionado por se tratar de um procedimento empírico, em que há uma investigação sistemática, rigorosa e com determinado nível de controle, semelhante a um experimento. Porém, se diferencia de um experimento, pois, à atribuição de tratamentos a sujeitos não pode ser baseada em aleatoriedade, mas em características dos sujeitos ou dos próprios objetos.

O experimento visa avaliar se a técnica BPPA é acurável e reprodutível para que seja possível identificar suas limitações. Dessa forma, objetivou-se avaliar a técnica análise de pontos de processos de negócio, com o propósito de avaliá-la em relação à reprodutibilidade e acurácia para identificar suas limitações, do ponto de vista do usuário, no contexto da mensuração de projetos de BPM para organizações.

Para a avaliar os resultados obtidos no experimento, dois critérios foram considerados: reprodutibilidade e acurácia. Segundo trabalhos relacionados à reprodutibilidade em APF, o coeficiente de variação esperado deve ser um valor inferior a 17,67% [19], [20], [21]. Esse valor foi usado como referência, pois o resultado pode ser mantido sem alterações substanciais ao longo dos anos para medições de técnicas de medição funcional. Para acurácia, a comparação entre os valores encontrados pelos sujeitos com o valor real deveria resultar em um coeficiente de variação inferior a 10,71%

1. O quase-experimento é denominado neste artigo como experimento

que é a média dos coeficientes de variação encontrados em trabalhos relacionados a acurácia para APF [22], [23], [24], [25], [26], [27], [28].

Como referência, para avaliar a acurácia e reprodutibilidade em termos de desvio padrão, foram usados valores de referência publicados na literatura sobre experimentos similares envolvendo APF. Valores de referência de APF foram usados devido à falta de experiências com técnicas para BPM. Deste modo, a autora assumiu que se tratam de técnicas equivalentes para domínios diferentes, podendo então utilizar referências de APF para BPPA [1].

Este artigo foi caracterizado da seguinte forma [29]:

- **Gênero “prático”**: conhecimento de pesquisa ligado à prática de conhecimento científico para fins explícitos de intervenção na realidade, sem perder o rigor metodológico, pois a partir da execução prática dos sujeitos participantes no cálculos dos Pontos de Processos de Negócio (BPPs<sup>2</sup>), a avaliação pode ser efetuada.
- **Abordagem “exploratória”**: abordagem que proporciona maior familiaridade com o problema, com objetivo de torná-lo explícito ou construir hipóteses. Este artigo objetiva verificar a hipótese de que a técnica BPPA é acurável e reprodutível.
- **Natureza “aplicada”**: gera conhecimentos para aplicação prática dirigida à solução de problemas específicos. A avaliação da técnica contribui para a geração de conhecimentos reprodutibilidade e acurácia.
- **Abordagem “Mista – qualitativa e quantitativa”**: quantitativa para ser capaz de mensurar os resultados obtidos com o procedimento estatístico e qualitativa, para que com base nos valores quantitativos obtidos, seja possível avaliar as limitações da técnica.
- **Procedimento “Quase-experimento”**: selecionado pois, flexibiliza as condições impostas pelos experimentos puros, os quais são usados para verificar o relacionamento entre a causa e o efeito de uma teoria. Neste artigo, com o objetivo de identificar se a técnica BPPA (causa) é acurável e reprodutível (efeito). Especificamente, um quase-experimento foi conduzido, pois, há a possibilidade de flexibilização do experimento, no qual a atribuição de tratamentos a sujeitos não pode basear-se na aleatoriedade, mas emerge das características dos sujeitos ou dos próprios objetos.

## 5. Protocolo do experimento

Esta seção descreve os resultados do planejamento do experimento, guiados pela definição de Wohlin [18], que define a necessidade de saber se o experimento é realmente capaz de avaliar o objeto de estudo de modo que seja possível realizar uma abordagem apropriada para responder à questão de pesquisa [18].

2. Do inglês: *Business Process Points*

**Seleção dos sujeitos**: foram selecionados por amostragem não probabilística por conveniência. Participaram oito alunos da disciplina “Gestão computacional de processos de negócio” do Programa de Pós-graduação acadêmica *stricto sensu* em Sistemas de Informação da Universidade A, 46 alunos da disciplina “Prática e Gerenciamento de Projetos” do quarto ano da graduação em Sistemas de Informação da Universidade A, e 2 alunos da disciplina “Tópicos Especiais em Computação DLXXXVII: Modelagem e Gerenciamento de Processos de Negócio” do programa de pós-graduação acadêmica *stricto sensu* em Informática da Universidade B.

Esta escolha se deve as limitações de tempo, custo e de acessibilidade para usar profissionais de medição de software, pois a disponibilidade de tempo desses profissionais é menor que a disponibilidade dos alunos. O tempo necessário para o treinamento sobre BPPA e da realização da contagem funcional dos modelos de processos de negócio, demandou aproximadamente quatro horas, o que torna a disponibilidade de voluntários para o experimento muito baixa. Esses profissionais demandam um custo relativamente alto, pois, as horas necessárias para a realização do experimento interfeririam nos seus horários de trabalho. E por fim, limitações de acessibilidade, pois, o acesso à quantidade de profissionais necessários para o experimento é muito mais difícil do que o acesso aos estudantes de graduação e de pós-graduação, devido aos fatores citados a cima.

A seleção dos sujeitos para o experimento evita a aleatoriedade, buscando propriedades e características inerentes a formação acadêmica dos sujeitos participantes. A grande maioria dos sujeitos selecionados já atua no mercado de trabalho tendo contato diário com processos de negócio.

**Seleção das variáveis**: com base em desempenho e focadas em identificar o quão bem os sujeitos são capazes de realizar o experimento. **Variáveis dependentes** – reprodutibilidade e acurácia; **variáveis independentes** – BPPA.

**Instrumentação**: inclui: materiais de treinamento, documentação sobre BPPA, objetos experimentais, formulário de caracterização de sujeito, conforme descrito abaixo:

- **Materiais de treinamento** – ao início do experimento os sujeitos foram expostos a um treinamento prévio, com informações sobre o processo de contagem funcional de BPPA em que as definições foram apresentadas. O material usado nesta apresentação, um poster de BPMN foi disponibilizado para fins de consulta das determinações da técnica;
- **Documentação BPPA** – os sujeitos participantes receberam também um relatório técnico sobre a técnica BPPA [13] para servir de material de apoio ao decorrer da contagem. Nesse relatório continha todas as informações pertinentes à definição de contagem funcional e as regras de sua aplicação.
- **Objetos experimentais** – três especificações de modelos de processo de negócio BPMN foram usados. Foi disponibilizado ainda, um documento-resposta, em que nos locais indicados os participantes deveriam informar as repostas dos passos das contagens e ao fim, informar o valor final em BPPs. O tempo

gasto no experimento também foi registrado.

- **Formulário de caracterização de sujeito** – todos os sujeitos responderam a um questionário de caracterização que continha uma breve descrição do experimento e informações de proteção à confidencialidade dos dados. Em poucas perguntas foi possível caracterizar a atuação dos sujeitos e seu conhecimento em medição funcional, processo de negócio e BPMN.

**Formulação das hipóteses:** A fim de avaliar se a técnica BPPA é reprodutível e acurável, duas hipóteses de pesquisa foram definidas, abordando as variáveis dependentes em relação a variável independente.

- **Reprodutibilidade:**
  - H0 – A técnica BPPA apresenta um coeficiente de variação de reprodutibilidade superior a 17,67%, ou seja, a técnica BPPA apresenta um grau de reprodutibilidade maior que o valor de referência.
  - H1 – A técnica BPPA apresenta um coeficiente de variação de reprodutibilidade igual ou inferior a 17,67%, ou seja, a técnica BPPA apresenta grau de reprodutibilidade equivalente ou menor ao valor de referência.
- **Acurácia:**
  - H0 – A técnica BPPA apresenta um coeficiente de variação de acurácia superior a 10,71%, ou seja, a técnica BPPA apresenta um grau de acurácia maior do que o valor de referência.
  - H1 – A técnica BPPA apresenta um coeficiente de variação de acurácia igual ou inferior a 10,71%, ou seja, a técnica BPPA apresenta um grau de acurácia equivalente ou menor que o valor de referência.

**Desenvolvimento do experimento:** O experimento foi realizado no final do segundo semestre letivo de 2017. A todos os sujeitos envolvidos foi oferecido um treinamento sobre a técnica BPPA, no qual conceitos, definições e a base de cálculo da técnica foram explicados. Após esse processo, cada um dos sujeitos recebeu o respectivo modelo de processo de negócio modelado em BPMN, em que a partir dos conhecimentos obtidos com a palestra introdutória, agregado a possibilidade de consultas aos materiais de apoio, os sujeitos participantes estimaram o tamanho funcional dos modelos de processo de negócio, desta forma, ao fim do experimento, obtiveram-se valores pertinentes aos três modelos.

Dois dos modelos de processo de negócio foram selecionados a partir do experimento realizado pela autora [2] da técnica, que usou os mesmos modelos de processo de negócio para avaliar a usabilidade da técnica quando ela foi desenvolvida. Já o terceiro modelo de processo de negócio foi selecionado pelos autores deste artigo, em que um modelo de processo de negócio maior e não necessariamente mais complexo foi selecionado.

Com base nos números de BPPs calculados pelos sujeitos, as avaliações referentes à reprodutibilidade e à acurácia foram realizadas. Os valores reais foram calculados previamente pelos pesquisadores proponentes, por serem os únicos especialistas em BPPA. A Tabela 2 apresenta os valores reais considerados. A definição desses valores foi definida da seguinte forma: para os modelos 1 e 2 os valores definidos pela autora [2] foram considerados e para o modelo 3, a média entre os dois autores deste artigo foi considerada.

Para fins da avaliação da reprodutibilidade, o desvio padrão do valor calculado por cada sujeito participante em relação à média dos valores encontrados pelos demais sujeitos e por fim, à média de todos os desvios padrões. O valor encontrado do coeficiente de variação foi comparado ao valor de referência 17,67%, em que para que seja classificado como reprodutível, a técnica deve obter um valor equivalente ou inferior a ele.

Para fins da avaliação da acurácia, foram verificados os desvios padrões entre os valores calculados por cada sujeito participante em relação ao valor real e ao fim, a média de todos os desvios padrões. O valor encontrado do coeficiente de variação foi comparado ao valor de referência de 10,71%, em que para que seja classificado como acurável, a técnica deve obter um valor equivalente ou inferior a ele.

A avaliação dos resultados obtidos com a realização do experimento foi realizada usando estatística descritiva, a fim de avaliar as hipóteses de pesquisa, objetivando obter informações o suficiente para identificar se a técnica BPPA é reprodutível e acurável e conseqüentemente suas limitações.

Para fins desse experimento o anonimato dos sujeitos foi mantido, com o objetivo de que houvesse um maior índice de comprometimento e veracidade no experimento. A fim de permitir resultados mais consistentes, os sujeitos participantes tiveram a oportunidade de realizar em um prazo de até duas semanas, para que o andamento de suas atividades acadêmicas, bem como as possíveis atividades profissionais executadas em paralelo não fossem prejudicadas.

## 6. Resultados obtidos

Com os 58 participantes do experimento foi possível obter um total de 137 respostas, divididas não-igualmente entre os três modelos de processo. Os resultados obtidos com a contagem dos BPPs estão divididos em três partes: os valores obtidos com o cálculo total de BPPs, os valores obtidos com o cálculo das funções de dados e os valores obtidos com o cálculo das funções de transações.

Durante o treinamento prévio sobre a técnica BPPA aos sujeitos participantes, uma avaliação comportamental foi realizada a fim de identificar fatores que pudessem impactar negativamente o resultado do quase-experimento. Foi possível observar que alguns sujeitos participantes estiveram em ociosidade, o que nos permite fazer avaliações dos resultados omitindo os seus valores. Dessa forma, dos 137 resultados obtidos, 30 foram continham a referência de ociosidade. Outros 23 experimentos foram alvo de avaliações devido a ocasionais atrasos de alguns sujeitos participantes o que abre precedente para a falta de informação suficiente para realizar

o experimento. Por fim, dos 84 experimentos restantes, 15 obtiveram valores extremos, o que pode impactar muito no cálculo do desvio padrão, e acarretar em valores negativos para toda a avaliação de dados. Para fins desse experimento, ambas as amostras foram consideradas.

### 6.1. Dados da reprodutibilidade

A Tabela 1 apresenta os resultados obtidos com a avaliação da reprodutibilidade. A primeira linha representa a média entre os resultados de todos os sujeitos participantes exceto dos que apresentaram indícios de ociosidade e atrasos. Na segunda linha o desvio padrão em relação à média é apresentado e por fim, o coeficiente de variação. Para o modelo de processos de negócio 1 o coeficiente de variação é de 34%, um valor 1,93 vezes maior que o valor de referência de 17,67%. Para os modelos de processo de negócio 2 e 3, 43,64% e 42,24% foram obtidos, respectivamente, valores 2,47 e 2,39 vezes maiores que o valor de referência.

Para o cálculo das funções de dados (uma etapa do cálculo total de BPPs) valores de coeficiente de variação de 34,39%, 48,90% e 41,66% foram encontrados para os modelos de processo de negócio 1, 2 e 3, respectivamente. Já nas funções de transações 59,67%, 84,73% e 65,62% foram os coeficientes de variação encontrados para os mesmos modelos. Valores em média 1,39, 2,23 e 1,77, respectivamente, maiores que os obtidos na estimativa funcional de dados.

A terceira linha, da Tabela 1, apresenta a média obtida entre os resultados de todos os sujeitos participantes exceto dos que apresentam indícios de ociosidade, atrasos e valores extremos. Aqui, os valores extremos foram desconsiderados para que análises relacionadas aos resultados sem esses valores, sejam capazes de identificar se esses valores tem significância relevante no cálculo da reprodutibilidade. Os valores obtidos na estimativa funcional do modelo de processo de negócio 1 obteve um coeficiente de variação de 30,92%, um valor 1,75 vezes maior que o valor de referência de 17,67%. Similar a isso, ocorreu com os modelos de processo de negócio 2 e 3, em que 35,04% e 37,56%, respectivamente, foram obtidos como coeficiente de variação resultando em 1,98 vezes e 2,12 vezes, respectivamente, maiores que o valor referência citado para o resultado total do cálculo de BPPs.

Em relação à estimativa funcional de funções de dados, coeficientes de variação de 36,29%, 52,34% e 34,30% foram obtidos para os modelos de processo de negócio 1, 2 e 3, respectivamente. E para a estimativa funcional de funções de transação, coeficientes de variação de 54,34%, 71,07% e 47,57% foram obtidos para os modelos de processo de negócio 1, 2 e 3, respectivamente. Novamente, valores cerca de 2,93, 1,26 e 0,34 vezes maiores que os referentes ao coeficiente de variação para funções de dados.

### 6.2. Dados da acurácia

Para avaliação da acurácia o coeficiente de variação é calculado com base no valor real estimado pelos pesquisadores proponentes. Para o modelo de processos de negócio

1, 53%, para o modelo 2, 70% e para o modelo 3, 53,50%, conforme é possível observar na primeira linha da Tabela 2.

A segunda linha da Tabela 2 representa o cálculo do desvio padrão com base no valor real obtido pelos sujeitos exceto os que apresentaram indícios de ociosidade e atrasos. Para o modelo 1 o coeficiente de variação é de 64,34% e para os modelos de processo negócio 2 e 3, 53,44% e 93,58% foram registrados. Esses são valores 6,05, 4,99 e 8,74 vezes maiores que o valor de referência de 10,71%.

Em relação às funções de dados coeficientes de variação de 66,61%, 76,06% e 98,20% foram calculados para os modelos 1, 2 e 3, respectivamente. Para a estimativa funcional de transações valores 79,89%, 73,52% e 111,98% foram obtidos no cálculo do coeficiente de variação para os mesmos modelos. Esses valores são em média 1,16 e 1,10 vezes maiores que os modelos 1 e 3 respectivamente. O modelo 2 teve valores inferiores nas funções de transações em relação as funções de dados de 0,3 vezes.

Com base no resultado da medição realizada pelos sujeitos, exceto os que apresentaram indícios de ociosidade, atrasos e valores extremos, para o modelo 1, foi obtido um coeficiente de variação de 63,81%, um valor 5,96 vezes acima do valor de referência. Assim como na avaliação de reprodutibilidade, ocorreram resultados para os resultados dos modelos 2 e 3, 43,36% e 60,15%, respectivamente, foram obtidos no coeficiente de variação. Em média valores 4,05 e 5,61 vezes maiores, respectivamente, do que o valor de referência para cada um dos modelos de processo.

Para o cálculo das funções de dados valores de coeficiente de variação de 63,61%, 112,13% e 95,31% foram obtidos respectivamente para os modelos 1, 2 e 3. Já em relação a funções de transações, os valores 38,48%, 146,81% e 174,76% foram obtidos. Esses valores são 3,24 e 7,62 vezes maiores nas funções de transações do que nas funções de dados para os modelos 2 e 3. Entretanto, o modelo 1, obteve valores 2,35 vezes inferior nas funções de transações em relação as funções de dados.

## 7. Avaliação dos Resultados

Esta seção discute os resultados obtidos no estudo empírico conduzido, baseando-se nos critérios pré-definidos para a avaliação da técnica. Limitações da técnica e as ameaças a validade do experimento também são discutidas.

Em resumo, o experimento realizado apresentou resultados que indicam que a técnica BPPA não é reprodutível nem acurável. Os valores obtidos são reflexo de dois problemas em potencial: a não realização de um experimento piloto e as falhas na definição dos processos referentes à contagem funcional de BPPs no relatório técnico.

Os resultados obtidos no coeficiente de variação para reprodutibilidade, são, em média, 2,5 vezes maiores que os valores de referência, tanto para os resultados que consideram valores extremos quanto para os que não consideram. A diferença, apesar de existir, não melhora substancialmente os resultados obtidos, permitindo que eles sejam reprodutíveis. Essa baixa diferença em relação às duas análises permite identificar que houve falhas e dificuldades na estimativa

Tabela 1. ANÁLISE DA REPRODUTIBILIDADE

Sem considerar		Total de BPPs			Funções de dados			Funções de transações		
		Mod.1	Mod.2	Mod.3	Mod.1	Mod.2	Mod.3	Mod.1	Mod.2	Mod.3
	Média	37,74	69,06	80,07	21,7	44,97	49,04	16,04	24,1	31,04
atrasos e ociosidade	Desvio padrão	12,83	30,14	33,74	7,44	21,99	20,43	9,57	20,42	20,36
	Coef. variação	34,00%	43,64%	42,14%	34,29%	48,90%	41,66%	59,67%	84,73%	65,62%
	Média	34,35	54,5	67,78	20,61	37,32	43,65	13,74	17,18	24,13
atrasos, ociosidade e valores extremos	Desv. padrão	10,62	19,1	25,46	7,48	19,53	14,97	7,47	12,21	11,48
	Coef. variação	30,92%	35,04%	37,56%	36,29%	52,34%	34,30%	54,34%	71,07%	47,57%
	Valor ref.:	<= 17,67%								

Tabela 2. ANÁLISE DA ACURÁCIA

Sem considerar		Total de BPPs			Funções de dados			Funções de transações		
		Mod.1	Mod.2	Mod.3	Mod.1	Mod.2	Mod.3	Mod.1	Mod.2	Mod.3
	Valor real	53	70	53,5	19	35	31,5	34	35	22
atrasos e ociosidade	Desvio padrão	34,1	37,41	50,07	12,66	26,62	30,93	27,16	25,73	24,64
	Coef. variação	64,34%	53,44%	93,58%	66,61%	76,06%	98,20%	79,89%	73,52%	111,98%
atrasos, ociosidade e valores extremos	Desvio padrão	33,82	30,36	32,18	12,09	39,25	30,02	13,08	51,38	38,45
	Coef. variação	63,81%	43,36%	60,15%	63,61%	112,13%	95,31%	38,48%	146,81%	174,76%
	Valor ref.:	<= 10,71%								

funcional com quase todos os sujeitos que participaram do experimento e os valores extremos indicam que essas falhas impactam muito nos resultados e que precisam ser evidenciadas.

Para os valores de acurácia a diferença em relação ao valor de referência é ainda maior, em média, 6,5 vezes a cima do esperado e novamente os resultados que consideram valores extremos não melhoram o coeficiente de variação obtido ao ponto de permitir que a técnica seja acurável. Esse valor evidencia a percepção de que assim como APF apesar de os passos serem seguidos nem sempre o valor de BPPs encontrado será próximo ao valor real, porém, para BPPA os valores obtidos são expressivamente ruins.

Entretanto, os resultados indicam que o maior índice de divergência ocorreu na estimativa funcional de transações, o que permite entender que os sujeitos tiveram maior dificuldade nesse passo devido a problemas não identificados previamente e dificuldades possivelmente técnicas em relação à definição da própria BPPA ou do experimento.

Se um experimento piloto fosse realizado previamente à aplicação do experimento, com uma amostra menor, potenciais problemas poderiam ter sido identificados e prevenidos em relação à execução do experimento e as falhas existentes no relatório técnico. Seria possível identificar que o treinamento oferecido aos sujeitos, assim como o relatório técnico oferecido como material de apoio, gerariam problemas de interpretação. Seria possível identificar também que o tempo permitido para a realização do experimento e para a flexibilização de local devido às atividades didáticas e profissionais dos sujeitos participantes, permitiriam potenciais distrações e falta de comprometimento com o experimento. Outro fator que deve ser considerado é que poderia ter sido possível ser mais assertivo em relação à amostra final do experimento e, por fim, uma abordagem específica poderia ter sido planejada para diminuir os índices de ociosidade dos sujeitos participantes.

## 7.1. Limitações da técnica BPPA

Fatores importantes foram identificados pelos pesquisadores proponentes que podem ter contribuído para resultados expressivamente ruins.

O relatório técnico que norteou o experimento foi desenvolvido em 2012 pela autora da técnica BPPA [13]. Esse documento define e descreve cada etapa da contagem funcional dos BPPs. Entretanto, foram identificados diversos pontos que geram interpretações diferentes em seu entendimento, o que dificulta a compreensão. Apesar de a técnica ser desenvolvida com base em APF, a adaptação para BPM deixou pontos a serem melhorados.

No passo de determinação da fronteira a identificação do participante principal do processo de negócio modelado em BPMN é realizada, de modo que esse participante, seja a representação gráfica da fronteira, através de piscinas ou raias. Aqui, ambiguidades na compreensão acontecem, pois, a percepção individual em relação ao participante principal do modelo de processo de negócio modelado em BPMN, pode variar de sujeito para sujeito.

No passo de identificação das funções de dados novamente problemas relacionados à percepção ocorrem, pois existem diversas ferramentas usadas para modelar um processo de negócio em BPMN e apesar de ser uma notação de modelagem de processo de negócio muito bem estabelecida, essas ferramentas tem suas características individuais. O relatório técnico desenvolvido pela autora, apresenta a representação gráfica de cada elemento de BPMN e sua nomenclatura, que nem sempre coincide com a representação gráfica de determinada ferramenta de modelagem. Se o modelo de processo de negócio não estiver modelado exatamente com as representações definidas no relatório técnico, problemas de compreensão podem ocorrer.

No decorrer do relatório técnico, problemas relacionados ao uso de termos específicos, sem definição prévia de seu significado no documento, foram alvos de ambiguidade na

compreensão. A ausência da definição desses termos, permite que diferentes pessoas compreendam de modo diferente informações que impactam fortemente no resultado da estimativa funcional. Termos como: classificação, identificação, determinação e agrupamento, foram empregados diversas vezes com diferentes intenções, o que novamente permite problemas na realização da estimativa.

No relatório técnico há a especificação de todos os elementos de BPMN, entretanto, a contagem pode ser impactada quando um elemento é interno a outro, como no caso de subprocessos expandidos. A identificação de estruturadas de dados e de atividades pode sofrer diferenças se forem considerados ou não os valores internos aos subprocessos. Não são encontradas informações sobre os subprocessos expandidos (por serem definidos como elementos de atividades), que determinam se devem ou não ser contados como atividades além de suas atividades e estruturas de dados internas. Em ambas as situações o valor encontrado pode impactar no cálculo final de BPPs.

Por fim, problemas na classificação dos tipos de funções de transações ocorreram. As funções de transações podem ser classificadas como Entrada Externa (EE), Saída Externa (SE) e Consulta e Externa (CE), entretanto, nem sempre os modelos de processos de negócio são compostos por mais de uma raia ou piscina para definir entradas ou saídas na fronteira. Assim, apenas as CEs sobram, mas também não podem ser usadas pois realizam consultas externas a fronteira, o que faz com que os valores transações sejam zerados, impactando no valor final de BPPs.

Esses fatores identificados como limitações da técnica BPPA, não indicam que a técnica seja ruim, eles apenas são fatores que indicam que o relatório técnico usado como base para a realização desse experimento deve ser melhorado nos pontos citados, de modo que, a técnica venha a obter melhores e maior compreensão, permitindo que possa ter melhores índices de reprodutibilidade e acurácia.

## 7.2. Ameaças à validade

Ameaças à validade de conclusão, interna, de construção e externa são discutidas a seguir a fim de maximizar a validade desse experimento.

**7.2.1. Ameaça a validade de conclusão.** A fim de assegurar a relação estatística entre tratamento e resultado deste experimento, técnicas estatísticas descritivas foram aplicadas incluindo média, desvio padrão e coeficiente de variação. Não foram aplicadas estatísticas inferenciais, o que pode indicar uma possível ameaça à validade. A heterogeneidade aleatória dos sujeitos, com baixa ou nenhuma experiência profissional com medição funcional e/ou BPMN e/ou processos de negócio também pode representar uma ameaça aos resultados, itens que podem ser abordados futuramente como foco de outras experiências.

**7.2.2. Ameaça a validade interna.** A fim de assegurar a relação entre causa e efeito entre tratamento e resultado, o mesmo material de treinamento sobre a técnica foi usado

em ambos os grupos experimentais, evitando conflito de informações e diferenciação. Entretanto, a diferença entre o perfil dos sujeitos participantes pode impactar os resultados desse experimento, pois, no grupo experimental sujeitos com experiência de 0 à 8 anos foram considerados. O fato do experimento não ter sido conduzido em grupo único, também é considerado uma ameaça à validade interna, pois, podem ter ocorrido diferentes tratamentos e reações a esses tratamentos em diferentes grupos. Os cálculos sobre acurácia e reprodutibilidade foram realizados três vezes, uma vez para cada modelo de processo de negócio considerado no experimento, como forma de diminuir as ameaças à validade interna e de construção do experimento.

**7.2.3. Ameaça à validade de construção.** A fim de assegurar que o tratamento reflete a causa e o resultado reflete o efeito, a chamada teoria da observação, as seguintes ações foram tomadas: um treinamento prévio sobre a técnica foi cedido aos sujeitos; um documento-resposta foi disponibilizado para que pudessem usá-lo sistematicamente para apresentar os resultados de forma o mais uniforme possível em relação aos demais participantes; as hipóteses do experimento não foram expostas aos sujeitos, a fim de evitar que criassem hipóteses sobre essas hipóteses; e para a relação de efeito as seguintes ações foram tomadas: as variáveis dependentes foram seguidas de acordo com o VIM [5] como feito no estudo de [30]. Ameaças foram atribuídas em decorrência do tempo permitido para a realização do experimento, esse fator seria uma ameaça à validade da construção, entretanto, uma vez que se trata de um quase-experimento, ele é minimizado, pois, diferente de um experimento comum ele não precisa ser totalmente controlado.

**7.2.4. Ameaça à validade externa.** A fim de assegurar que a generalização dos resultados do experimento seja possível, foram evidenciados fatores que podem impedir sua reprodução: o uso e valores de referência baseados em autores de estudos relacionado à APF é considerado como ameaça à validade do quase-experimento, pois processos de negócios e pontos funções possuem níveis de abstrações diferentes. Considerou-se também como ameaças à validade o local onde experimento ocorreu, os tipos de sujeitos e o treinamento prévio realizado, considerados fatores que poderiam impedir que a generalização dos resultados ocorresse de forma representativa. Entretanto, existe a consciência de que estudos empíricos ainda são necessários com outros grupos experimentais e em ambiente controlado para que a validade externa seja assegurada. A ausência de profissionais em medição funcional poderia ser caracterizada com uma ameaça externa à validade, porém, o seu efeito pôde ser mitigado por ser conduzido com alunos de disciplinas relacionadas a gestão de projetos, proporcionando boa validade interna à custa da validade externa dos resultados.

## 8. Conclusões

A avaliação da técnica BPPA foi realizada com 58 alunos de graduação e pós-graduação como sujeitos a partir de

um quase-experimento que obteve inicialmente 137 respostas. Nesse experimento os sujeitos realizaram a estimativa funcional de três modelos de processo de negócio usando a técnica BPPA e estimando a quantidade de BPPs para cada modelo. A análise dos resultados da técnica apresentou valores substancialmente ruins referentes à reprodutibilidade e valores ainda piores em relação à acurácia. Essa ocorrência foi atribuída à ausência da execução prévia de um experimento piloto que auxiliaria na identificação de potenciais problemas de execução do experimento e de interpretação dos sujeitos, de modo que, erros e falhas potenciais na execução do experimento final poderiam ter sido evitados.

Problemas relacionados às limitações do relatório técnico que descreve a técnica BPPA foram discutidos e considerados fatores importantes no baixo índice de reprodutibilidade e acurácia da técnica. Esses fatores permitiram problemas de interpretação em relação aos passos necessários para a estimativa funcional dos BPPs. É provável que o resultado do experimento pudesse ser diferente se ambos os fatores tivessem sido melhor planejados.

De modo geral, trabalhos futuros podem ser direcionados para estudos que objetivem melhorar a formulação da técnica BPPA e suas etapas de medição, visando mitigar problemas de interpretação e ambiguidade na estimativa funcional e consequentemente aumentar os índices de reprodutibilidade e acurácia da técnica. Essas são ações necessárias para permitir que a técnica venha a ser disponibilizada para a indústria e permitir o uso a gerentes de processos de negócio em cenários de automação de processos de negócio.

## Referências

- [1] M. Baklizky, M. Fantinato, L. H. Thom, V. Sun, and P. C. Hung, “Business process point analysis: survey experiments,” *Business Process Management Journal*, vol. 23, no. 2, pp. 399–424, 2017.
- [2] M. Baklizky, M. Fantinato, L. H. Thom, V. Sun, E. P. Prado, and P. Hung, “Business process points—a proposal to measure bpm projects,” in *ECIS*, 2013, p. 2.
- [3] C. Monsalve, A. Abran, and A. April, “Measuring software functional size from business process models,” *International Journal of Software Engineering and Knowledge Engineering*, vol. 21, no. 03, pp. 311–338, 2011.
- [4] G. Muketha, “A survey of business processes complexity metrics,” 2010.
- [5] I. BIPM, I. IFCC, I. IUPAC, and O. ISO, “The international vocabulary of metrology—basic and general concepts and associated terms (vim),” *Joint Committee for Guides in Metrology*, vol. 3, 2012.
- [6] M. Dumas, M. La Rosa, J. Mendling, H. A. Reijers *et al.*, *Fundamentals of business process management*. Springer, 2013, vol. 1.
- [7] M. Weske, *Business process management: concepts, languages, architectures*. Springer Publishing Company, Incorporated, 2010.
- [8] M. Fantinato, I. M. S. Gimenes, and M. B. F. Toledo, *Product line in the business process management domain*. Boston, MA, USA, Auerbach Publications, 2010, pp. 497–530.
- [9] M. P. Barcellos, “Uma estratégia para medição de software e avaliação de bases de medidas para controle estatístico de processos de software em organizações de alta maturidade,” Ph.D. dissertation, Universidade Federal do Rio de Janeiro, 2009.
- [10] N. Fenton and J. Bieman, *Software metrics: a rigorous and practical approach*. CRC Press, 2014.
- [11] L. Lazic and N. Mastorakis, “Cost effective software test metrics,” *WSEAS Transactions on Computers*, vol. 7, no. 6, pp. 599–619, 2008.
- [12] J. Cardoso, “Business process control-flow complexity: Metric, evaluation, and validation,” *International Journal of Web Services Research*, vol. 5, no. 2, p. 49, 2008.
- [13] M. Baklizky and M. Fantinato, “Relatório técnico ppgsi-003/2012 fpa4bpm—function point analysis for business process management (v. 1.0),” 2012.
- [14] B. P. Model, “Notation (bpmm) version 2.0,” *OMG Specification, Object Management Group*, 2011.
- [15] T. Chandrakumar and S. Parthasarathy, “An approach to estimate the size of erp package using package points,” *Computer Standards & Interfaces*, vol. 47, pp. 100–107, 2016.
- [16] S. Abrahão, E. Mendes, J. Gomez, and E. Insfran, “A model-driven measurement procedure for sizing web applications: design, automation and validation,” in *Proceedings of the International Conference on Model Driven Engineering Languages and Systems*. Springer, 2007, pp. 467–481.
- [17] S. Abrahão, G. Poels, and O. Pastor, “A functional size measurement method for object-oriented conceptual schemas: design and evaluation issues,” *Software & Systems Modeling*, vol. 5, no. 1, pp. 48–71, 2006.
- [18] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén, *Experimentation in software engineering*. Springer Science & Business Media, 2012.
- [19] M. J. Connolly, “An empirical study of function points analysis reliability,” Ph.D. dissertation, Massachusetts Institute of Technology, Sloan School of Management, junho 1990.
- [20] G. C. Low and D. R. Jeffery, “Function points in the estimation and evaluation of the software process,” *IEEE Transactions on Software Engineering*, vol. 16, no. 1, pp. 64–71, 1990.
- [21] C. F. Kemerer, “Reliability of function points measurement: a field experiment,” *Communications of the ACM*, vol. 36, no. 2, pp. 85–97, 1993.
- [22] S. Abrahão and E. Insfran, “A metamodeling approach to estimate software size from requirements specifications,” in *Proceedings of the 34th Euromicro Conference on Software Engineering and Advanced Applications*. IEEE, 2008, pp. 465–475.
- [23] N. A. Adem and Z. M. Kasirun, “Automating function points analysis based on functional and non functional requirements text,” in *Proceedings of the 2nd International Conference on Computer and Automation Engineering*. IEEE, 2010, pp. 664–669.
- [24] T. Edagawa, T. Akaike, Y. Higo, S. Kusumoto, S. Hanabusa, and T. Shibamoto, “Function point measurement from web application source code based on screen transitions and database accesses,” *Journal of Systems and Software*, vol. 84, no. 6, pp. 976–984, 2011.
- [25] P. Fraternali, M. Tisi, and A. Bongio, “Automating function point analysis with model driven development,” in *Proceedings of the 2006 conference of the Center for Advanced Studies on Collaborative research*. ACM, 2006, p. 18.
- [26] E. Lamma, P. Mello, and F. Riguzzi, “A system for measuring function points from an er-dfd specification,” *The Computer Journal*, vol. 47, no. 3, pp. 358–372, 2004.
- [27] T. Miyawaki, J. Iijima, and S. Ho, “Measuring function points from vdm-sl specifications,” in *Proceedings of the International Conference on Service Systems and Service Management*. IEEE, 2008, pp. 1–6.
- [28] T. Uemura, S. Kusumoto, and K. Inoue, “Function-point analysis using design specifications based on the unified modelling language,” *Journal of software maintenance and evolution: Research and practice*, vol. 13, no. 4, pp. 223–243, 2001.
- [29] G. H. Travassos, D. Gurov, and E. Amaral, *Introdução à engenharia de software experimental*. UFRJ, 2002.
- [30] S. Abrahão, G. Poels, O. Pastor *et al.*, “Comparative evaluation of functional size measurement methods: An experimental analysis,” Ghent University, Faculty of Economics and Business Administration, Tech. Rep., 2004.