



Ana Raquel Ribeiro Henriques

Licenciada em Engenharia informática

***Framework* de avaliação da complexidade e
facilidade de compreensão de modelos BPMN**

Dissertação para obtenção do Grau de Mestre em
Engenharia Informática

Orientadores: Prof. Dr. João Araujo, Prof. Auxiliar,
Universidade Nova de Lisboa
Prof. Dr. Miguel Goulão, Prof. Auxiliar,
Universidade Nova de Lisboa



FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Setembro, 2018

***Framework* de avaliação da complexidade e facilidade de compreensão de modelos BPMN**

Copyright © Ana Raquel Ribeiro Henriques, Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa.

A Faculdade de Ciências e Tecnologia e a Universidade NOVA de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

Em memória dos meus Avós

AGRADECIMENTOS

Antes de mais, quero agradecer aos meus orientadores por demonstrarem vontade de trabalhar comigo e por me acompanharem no meu processo evolutivo ao longo desta fase académica na minha vida. Quero agradecer ainda por me darem autonomia e guiarem-me na direção correta através de críticas construtivas e sugestões.

À Ana Catarina por me ajudar na validação e dar conselhos sobre construção dos testes com participantes, assim como companhia durante a fase de testes e pela paciência pelos incómodos causados com a entrada e saída constante com os participantes dos testes.

Ao professor Dénis Silveira da Universidade Federal de Pernambuco no Brasil, por disponibilizar material profissional e realizar testes com colegas profissionais e os seus alunos.

A todos os participantes nos meus testes com o *eye-tracker*, que despenderam do seu tempo em fases críticas do ano letivo para me ajudarem a recolher os dados necessários para concluir o meu trabalho.

Ao professor Paulo Lopes, por me ajudar com a disponibilização dos serviços e disponibilidade de ajuda para qualquer dúvida que surgisse.

Aos meus amigos e colegas de curso, pela paciência e apoio nos momentos críticos do percurso universitário.

Por último e não menos importante, quero agradecer os meus pais por me apoiarem ao longo de todo o meu percurso escolar, não só emocional, mas sim com toda a ajuda necessária ao longo de todos os meus anos de vida.

RESUMO

A modelação dos processos de negócios é cada vez mais utilizada para definir e descrever os processos de negócio das organizações. Esta metodologia é utilizada pelos *stakeholders* dos processos. Esta deve-se ao facto destes modelos serem uma ponte entre os seus utilizadores e os seus modeladores.

Uma vez que esta modelação pode envolver várias pessoas, é natural que possam ocorrer problemas na construção e interpretação de modelos de processos de negócios. Estes problemas podem resultar quer da complexidade essencial da linguagem de modelação (p.e. *Business Process Modelling Notation* - BPMN), quer da complexidade accidental dos modelos produzidos. Uma maneira de perceber os defeitos existentes nos modelos, consiste na avaliação da complexidade e a consequente da facilidade na compreensão destes. Atualmente existem estudos sobre a qualidade dos modelos, referentes à complexidade e correção destes, mas estes estudos baseiam-se unicamente nos modelos, não recorrendo a dados externos (p.e. interpretação de voluntários) para perceber a sua compreensão.

O foco deste trabalho é avaliar a complexidade e a facilidade de compreensão dos modelos BPMN. Para proceder à avaliação será utilizada como ponto de partida a abordagem *Goal-Question-Metric*. Esta abordagem será utilizada para determinar os objetivos da avaliação e as questões que permitam responder aos objetivos, que por sua vez serão respondidas com a implementação de métricas, podendo estas serem novas ou já existentes. As métricas serão ainda alvo de uma validação teórica segundo as propriedades de *Weyuker*. A recolha de dados através das métricas, será realizada de duas maneiras complementares, o primeiro caso consiste em recolher informações estatísticas dos modelos, no segundo caso foram recolhidos de dados relativos à compreensão dos leitores, com auxílio do dispositivo *eye-tracker*. Este dispositivo permite recolher dados de determinadas reações faciais das pessoas, tais como, o movimento dos olhos, a dilatação das pupilas e a taxa de pestanejos num determinado intervalo de tempo.

A recolha de dados permitiu iniciar a análise da compreensão dos modelos ao nível dos seus elementos do modelo. De modo a proceder esta recolha de dados e análise, foi implementado um *framework*, que permitirá continuar a recolher mais dados e consequentemente obter novas análises e conclusões sobre os modelos BPMN.

Palavras-chave: Sistemas de informação, Modelação dos processos de negócio, BPMN, Atributos de Qualidade, Métricas de Complexidade e Compreensão, *Eye-Tracker*

ABSTRACT

Nowadays, business process modeling is increasingly used to define and describe the business processes of organizations. These modeling are not only used by the stakeholders of process. This makes business process models a common bridge among stakeholders with different backgrounds.

Since this modeling may involve several people, it rises some problems on construction and interpretation of the business process models. These problems may result either the essential complexity of the modeling language (e.g. Business Process Modelling Notation - BPMN) or from the accidental complexity of the models produced. One way to understand the existing defects on models, consists in evaluating the complexity and the consequent facility on comprehension of them. Currently, there are studies about the quality of models, referring to the complexity and correctness of the models, but these studies are based only on models, not resorting external data (e.g. interpretation of voluntaries) to understand your comprehension.

The focus of this work is addressed on the comprehension of BPMN models. We will use the Goal-Question-Metric approach to drive the evaluation. This approach will be used to determine the evaluation objectives and the questions that allow us to assess those objectives, which in turn will be answered with the data collected through metrics, which may be new or not. The metrics will be a target of a theoretical evaluation according the properties of Weyuker. The collection of data through metrics, will be done in two complementary ways: the first case on the collection of information solely from the model, the second case data were collected related to reader comprehension, with the help of an eye-tracker device. This device allows the collection of data from certain facial reactions of people, such as, eye movements, pupil dilatation and the blink rate in a certain time interval.

The data collection allowed to begin the analysis of the understanding of the models at the level of its elements of the model. In order to carry out this data collection and analysis, a framework was implemented, which will allow to continue collecting more data and consequently to obtain new analyzes and conclusions about BPMN models.

Keywords: Information Systems, Business Process Modeling, BPMN, Attributes of Quality, Complexity and Comprehension Metrics, Eye-tracker

ÍNDICE

Lista de Figuras	xvii
Lista de Tabelas	xxi
Acrónimos	xxv
1 Introdução	1
1.1 Contexto e Motivação	1
1.2 Objetivos do trabalho	3
1.3 Principais Contribuições Previstas	4
1.4 Estrutura do Documento	4
2 Enquadramento	7
2.1 Modelação de processos de negócio	7
2.2 BPMN	8
2.2.1 Outras notações	11
2.3 Sistema de Informação	12
2.3.1 Engenharia de Software	12
2.4 Controlo de Qualidade de software	13
2.4.1 Atributos de qualidade	13
2.4.2 GQM	14
2.5 Especificação de Métricas	15
2.5.1 OCL	15
2.5.2 Complexidade	16
2.5.3 Propriedades e categorização das métricas	18
2.6 Eye-tracking	19
2.6.1 Recolha e Análise de Dados	20
2.6.2 Limitações	20
2.7 Sumário	21
3 Trabalho Relacionado	23
3.1 Análise da qualidade dos modelos de processos de negócio	23
3.2 Análise cognitiva recorrendo ao Eye-tracking	27

3.3	Sumário	28
4	Métricas de avaliação dos Modelos BPMN	29
4.1	Diagrama GQM	29
4.1.1	Objetivo - Nível Conceptual	30
4.1.2	Questões - Nível Operacional	30
4.1.3	Métricas - Nível Quantitativo	32
4.2	Meta-Modelo BPMN	34
4.3	Definição de métricas propostas	37
4.4	Sumário	50
5	Implementação do Protótipo	51
5.1	Opções de Implementação para recolha de métricas	51
5.2	Tecnologias utilizadas	52
5.3	Funcionalidades do protótipo	53
5.4	Sumário	54
6	Validação Teórica	55
6.1	Validação	55
6.1.1	Métricas dos Modelos	56
6.1.2	Métricas do Eye-tracker	58
6.2	Sumário	61
7	Validação Empírica	63
7.1	Plano de experiências	63
7.2	Sumário	73
8	Análise dos Dados Recolhidos	75
8.1	Participantes	75
8.2	Análises Estatísticas	77
8.2.1	Duração	78
8.2.2	Correção	99
8.2.3	Correção vs Duração	105
8.2.4	NASA-TLX	106
8.3	Hipóteses	112
8.3.1	$H_{confianca}$	112
8.3.2	$H_{prestacao}$	114
8.3.3	$H_{Pparticipantes}$	114
8.3.4	$H_{EfeitoAprendizagem}$	115
8.4	Discussão dos Resultados	115
8.4.1	Ameaças à validade	116
8.4.2	Inferências	116

8.5	Sumário	117
9	Conclusões	119
9.1	Síntese do trabalho realizado	119
9.2	Contribuições	120
9.3	Limitações	120
9.4	Trabalho Futuro	120
	Bibliografia	123
A	Exemplos dos Modelos BPM	131
A.1	Diagramas de atividades	131
A.2	SPEM	132
A.3	BPEL	133
A.4	BPT	134
A.5	Nintex Workflow	135
B	Exemplo OCL	137
C	Modelos Avaliados	141
D	Questionários dos Testes com Voluntários	153
D.1	Grupo A	153
D.2	Grupo B	154
D.3	Grupo C	155
D.4	Usabilidade	156
E	Carta de Consentimento	159
F	Questionário Demográfico	161
G	Manual de Utilizador	163
G.1	Login na plataforma	163
G.2	Registo na plataforma	165
G.3	Submissão de um modelo no formato XPDL	166
G.4	Realização dos testes para a validação empírica	169
I	Meta-Modelo BPMN	173
II	Simbologia modelos BPMN	175
III	Métricas Standard	183
IV	Métricas Standard Eye-tracker	187

LISTA DE FIGURAS

2.1	Exemplo de um modelo BPMN retirado de [Inc17]	11
2.2	Modelo de abordagem GQM	15
2.3	Exemplo de um modelo <i>eye-tracker</i> Intrusivo	19
2.4	Exemplo de um modelo <i>eye-tracker</i> não intrusivo	19
4.1	Modelo de abordagem GQM proposto	30
4.2	Raiz do meta-modelo <i>Business Process Model and Notation</i> (BPMN) adaptado de [Cor14]	35
4.3	Excerto do meta-modelo BPMN correspondente à classe <i>FlowElementsContainer</i> adaptado de [Cor14]	35
4.4	Excerto do meta-modelo BPMN correspondente à classe <i>FlowElement</i> adaptado de [Cor14]	36
4.5	Excerto do meta-modelo BPMN correspondente à classe <i>MessageFlow</i> adaptado de [Cor14]	36
4.6	Excerto do meta-modelo BPMN correspondente à classe <i>ItemAwareElement</i> adaptado de [Cor14]	37
4.7	Excerto do meta-modelo BPMN correspondente à classe <i>Artifacts</i> adaptado de [Cor14]	37
5.1	Esquema da arquitetura do protótipo	53
8.1	<i>BoxPlot</i> dos tempos dos participantes por grupos	79
8.2	<i>BoxPlot</i> dos tempos dos participantes por ordem do grupo A	80
8.3	<i>BoxPlot</i> dos tempos dos participantes por pergunta do grupo A	82
8.4	<i>BoxPlot</i> dos tempos dos participantes por posição do grupo A	83
8.5	<i>BoxPlot</i> dos tempos dos participantes por dificuldade do grupo A	84
8.6	<i>BoxPlot</i> dos tempos dos participantes por ordem do grupo B	85
8.7	<i>BoxPlot</i> dos tempos dos participantes por pergunta do grupo B	86
8.8	<i>BoxPlot</i> dos tempos dos participantes por posição do grupo B	87
8.9	<i>BoxPlot</i> dos tempos dos participantes por dificuldade do grupo B	88
8.10	<i>BoxPlot</i> dos tempos dos participantes por ordem do grupo C	89
8.11	<i>BoxPlot</i> dos tempos dos participantes por pergunta do grupo C	90
8.12	<i>BoxPlot</i> dos tempos dos participantes por posição do grupo C	91

8.13	<i>BoxPlot</i> dos tempos dos participantes por dificuldade do grupo C	92
8.14	<i>BarChart</i> das respostas dos participantes do grupo A	100
8.15	<i>BarChart</i> das respostas dos participantes do grupo B	101
8.16	<i>BarChart</i> das respostas dos participantes do grupo C	103
A.1	Exemplo de um diagrama de atividades [AN05]	131
A.2	Exemplo de um diagrama SPEM [Spe]	132
A.3	Exemplo de um diagrama BPEL [Cened]	133
A.4	Exemplo de um diagrama BPT [Fered]	134
A.5	Exemplo de um diagrama <i>Nintex workflow</i> [Nin]	135
B.1	Exemplo de um diagrama de classes [RG02]	137
C.1	Modelo fácil do grupo A	142
C.2	Modelo intermédio do grupo A	143
C.3	Modelo difícil do grupo A	144
C.4	Modelo fácil do grupo B	145
C.5	Modelo intermédio do grupo B	146
C.6	Modelo difícil do grupo B	147
C.7	Modelo fácil do grupo C	148
C.8	Modelo intermédio do grupo C	149
C.9	Modelo difícil do grupo C	150
C.10	Legenda dos modelos	151
G.1	Página principal da plataforma	163
G.2	Localização da funcionalidade de <i>Login</i> da plataforma	164
G.3	<i>Pop-up</i> de <i>Login</i>	164
G.4	Página de utilizador <i>Logado</i> na plataforma	165
G.5	<i>Pop-up</i> do registo na plataforma	165
G.6	Página de submissão de modelos BPMN para análise	166
G.7	Escolha de ficheiros a submeter na plataforma	167
G.8	Apresentação de modelo escolhido para avaliação na plataforma	167
G.9	Página de apresentação dos resultados em <i>Loading</i>	168
G.10	Página com apresentação dos resultados da aplicação das métricas no modelo submetido	168
G.11	Página pessoal de um utilizador da plataforma	169
G.12	<i>Pop-up</i> de seleção de opções de teste	169
G.13	Carta de consentimento de realização e recolha de dados no teste	170
G.14	<i>Layout</i> de apresentação de modelo dos testes	171
G.15	<i>Layout</i> de apresentação do teste de usabilidade do teste	171
G.16	Questionário demográfico dos testes	172

I.1 Meta-modelo BPMN 174

LISTA DE TABELAS

4.1	Conjunto de métricas para a avaliação da compreensão dos modelos BPMN .	32
4.2	Métrica M2	38
4.3	Métrica M4	38
4.4	Métrica M5	39
4.5	Métrica M6	40
4.6	Métrica M7	40
4.7	Métrica M10	41
4.8	Métrica M15	41
4.9	Métrica M16	42
4.10	Métrica M17	43
4.11	Métrica M18	43
4.12	Métrica M19	44
4.13	Métrica M23	44
4.14	Métrica M24	45
4.15	Métrica M25	45
4.16	Métrica M33	46
4.17	Métrica M34	46
4.18	Métrica M32	47
4.19	Métrica M27	48
4.20	Métrica M28	48
4.21	Métrica M29	49
4.22	Métrica M30	49
4.23	Métrica M31	50
6.1	Propriedades de <i>Weyuker</i>	55
6.2	Avaliação das Métricas dos Modelos	57
6.3	Avaliação das Métricas do Eye-tracker	59
7.1	Ordens dos níveis de dificuldade dos modelos nos grupos de testes	65
7.2	Número de voluntários sem conhecimento expressivo	66
7.3	Número de voluntários com conhecimento expressivo	66
8.1	Número de participantes	76

8.2	Estatísticas descritivas dos participantes sem conhecimento expressivo (informação de resposta aberta)	76
8.3	Estatísticas descritivas dos participantes sem conhecimento expressivo (informação de resposta fechada)	76
8.4	Duração de todos os testes	78
8.5	Estatísticas dos tempos do grupo A	79
8.6	Estatísticas dos tempos das perguntas do grupo A	81
8.7	Estatísticas dos tempos das perguntas dos modelos do grupo A	81
8.8	Estatísticas descritivas dos tempos do grupo B	84
8.9	Estatísticas dos tempos das perguntas do grupo B	86
8.10	Estatísticas dos tempos das perguntas dos modelos do grupo B	86
8.11	Estatísticas descritivas dos tempos do grupo C	88
8.12	Estatísticas dos tempos das perguntas do grupo C	90
8.13	Estatísticas dos tempos das perguntas dos modelos do grupo C	90
8.14	Estatísticas descritivas dos tempos dos diferentes grupos	92
8.15	Estatísticas descritivas dos tempos dos diferentes grupos por posição	93
8.16	Estatísticas descritivas dos tempos dos diferentes grupos por dificuldade	96
8.17	Percentagem de respostas corretas do grupo A por pergunta	99
8.18	Percentagem de respostas corretas do grupo B por pergunta	101
8.19	Percentagem de respostas corretas do grupo c por pergunta	103
8.20	Legenda testes NASA-TLX	106
8.21	A exigência mental para este teste foi a adequada?	107
8.22	A exigência física para este teste foi a adequada?	107
8.23	A exigência temporal para este teste foi a adequada?	107
8.24	Como considera o seu esforço aplicado durante o teste?	107
8.25	Qual o seu nível de frustração durante o teste?	108
8.26	Como classificaria a sua performance no teste?	108
8.27	A exigência mental para este teste foi a adequada?	109
8.28	A exigência física para este teste foi a adequada?	109
8.29	A exigência temporal para este teste foi a adequada?	109
8.30	Como considera o seu esforço aplicado durante o teste?	109
8.31	Qual o seu nível de frustração durante o teste?	110
8.32	Como classificaria a sua performance no teste?	110
8.33	A exigência mental para este teste foi a adequada?	110
8.34	A exigência física para este teste foi a adequada?	110
8.35	A exigência temporal para este teste foi a adequada?	111
8.36	Como considera o seu esforço aplicado durante o teste?	111
8.37	Qual o seu nível de frustração durante o teste?	111
8.38	Como classificaria a sua performance no teste?	111
8.39	Testes com aumento da taxa de sucesso e os seus respetivos dados performance NASA-TLX	112

8.40	Testes com diminuição da taxa de sucesso e os seus respetivos dados de frustração NASA-TLX	113
8.41	Perguntas relacionadas com os participantes de um modelo	114
III.1	Métricas de gráficos de acordo com o seu domínio	183
III.2	Métricas sobre a notação gráfica de acordo com o seu domínio	184
III.3	Métricas Petri Nets de acordo com o seu domínio	185
III.4	Métricas da complexidade de <i>software</i> de acordo com o seu domínio	185
III.5	Métricas da capacidade cognitiva de acordo com o seu domínio	185
III.6	Métricas de Shannon de acordo com o seu domínio	186
III.7	Métricas numéricas de acordo com o seu domínio	186
IV.1	Métricas relacionadas com o número de fixações	187
IV.2	Métricas relacionadas com os movimentos sacadas	188
IV.3	Métricas relacionadas com a duração das fixações	188
IV.4	Métricas relacionadas com os movimentos <i>scanpath</i>	189

ACRÓNIMOS

AOI	<i>Area of Interest.</i>
BPEL	<i>Business Process Execution Language.</i>
BPI	<i>Business Process Incubator.</i>
BPM	<i>Business Process Modeling.</i>
BPMN	<i>Business Process Model and Notation.</i>
BPT	<i>Business Process Technology.</i>
ES	<i>Engenharia de software.</i>
GQM	<i>Goal Question Metric.</i>
IEEE	<i>Instituto de Engenheiros Eletricistas e Eletrônicos.</i>
OCL	<i>Object Constraint Language.</i>
OMG	<i>Object Management Group.</i>
RR	<i>Regiões Relevantes.</i>
SI	<i>Sistemas de Informação.</i>
SPEM	<i>Software Process Engineering Metamodel.</i>
SPF	<i>Scan Path f-measure.</i>
SPP	<i>Scan Path Precision.</i>
SPR	<i>Scan Path Recall.</i>
UML	<i>Unified Modeling Language.</i>
XML	<i>eXtensible Markup Language.</i>
XPDL	<i>XML Process Definition Language.</i>

INTRODUÇÃO

Nesta secção é apresentada de maneira sucinta o contexto e o trabalho a desenvolver no âmbito desta dissertação, assim como as motivações por detrás da sua realização. É ainda apresentada uma breve descrição da estrutura deste documento.

1.1 Contexto e Motivação

Os **Sistemas de Informação (SI)** são desenvolvidos com as técnicas da **Engenharia de software (ES)** e são usados nos processos de negócio. Estes podem ser definidos como sendo uma combinação de pessoas, *hardware*, *software*, redes de comunicação, dados e políticas e procedimentos de armazenamento, acesso, transformação e disseminação de dados dentro de uma organização [OM08].

Os modelos de processos de negócio têm como objetivo modelar os processos de negócio das empresas/organizações de uma forma rigorosa [Wes10]. Além disto, os processos de negócio ajudam na elicitação de requisitos, da área da Engenharia de Requisitos, ajudando também a ES. Sem esquecer, que os SI fazem parte dos modelos dos processos de negócio, uma vez que estes representam todo o fluxo de dados dos processos. A área da Engenharia de Requisitos é aquela que define as necessidades e funcionalidades de um determinado produto ou serviço [Som16].

Os modelos de processos de negócios permitem ainda, especificar as atividades dos processos, quais os seus intervenientes e as informações utilizadas ou geradas pelas diversas atividades ou pelos intervenientes dos processos. A especificação de todas estas informações, permitem simular os processos de negócio e gerar os processos de uma maneira semiautomática. Uma outra característica dos modelos é a conexão entre os modeladores/programadores e os *stakeholders*, uma vez que à partida trata-se de uma linguagem perceptível para ambas as partes [Wes10].

A Engenharia de *Software* é aquela que suporta o desenvolvimento dos *SI*, de forma a que sejam bem construídos e fáceis de manter e evoluir. Estes sistemas são cada vez mais importantes devido ao crescimento do interesse dos setores económicos e empresariais em produtos de *software* [Bra05]. De um modo mais detalhado, a *ES* é a área encarregue de todas as etapas de desenvolvimento de *software*. Mais concretamente, o desenvolvimento começa na especificação dos requisitos e funcionalidades com os *stakeholders*, passando pela implementação e verificação de que os requisitos foram todos alcançados. Esta ainda contempla a manutenção e o desenvolvimento após o *software* ser entregue [Som16].

Como estes modelos são utilizados por um leque de profissionais bastante diverso é preciso ter em atenção, que a modelação dos processos pode não estar correta ou não ser perceptível para algumas dessas pessoas.

A dificuldade de modelação ou interpretação pode resultar de problemas da interpretação dos requisitos a modelar, da complexidade das linguagens de modelação a usar, ou da falta de especialização dos modeladores nessas linguagens. Se o modelador não compreender bem o significado ou enquadramento dos elementos de modelação ao seu dispor, pode vir a optar por soluções desnecessariamente complexas e, conseqüentemente, ainda menos compreensíveis para os restantes utilizadores [PM13; Pin+12].

Quanto maior a complexidade de um modelo, maior será a probabilidade de existirem defeitos na sua modelação. A existência de defeitos resulta num aumento da capacidade de compreensão exigida aos seus leitores e, conseqüentemente, leva a uma maior probabilidade dos leitores interpretarem os modelos de uma maneira incorreta.

A complexidade pode ser introduzida nos modelos de uma maneira acidental ou pode ser essencial. A complexidade acidental, pode gerar problemas nos modelos, tanto ao nível de introdução de erros e defeitos de modelação, assim como, na criação de ambigüidade [BK87]. Este género de complexidade pode ser atenuada, através da utilização de *frameworks* de avaliação da qualidade dos modelos.

Por outro lado, a complexidade essencial é necessária e é intrínseca ao problema alvo de modelação, por outras palavras, pode ocorrer quando o que se pretende modular possui grandes dimensões e diversos elementos a serem especificados [BK87].

É neste âmbito que se foca esta dissertação, ou seja, esta centra-se na qualidade dos modelos de processos de negócios, mais concretamente nos modelos *BPMN* [OMG11], do ponto de vista dos seus utilizadores. Os modelos podem ser avaliados segundo os atributos de qualidade de *software*, sendo estes: o seu funcionamento, fiabilidade, desempenho, usabilidade, segurança, compatibilidade, capacidade de manutenção e portabilidade [Jun+04]. Nesta dissertação, será dada uma maior ênfase à usabilidade destes, em particular, à facilidade de compreensão dos modelos por parte dos utilizadores.

Os modelos *BPMN* devem ser avaliados, pois, cada vez mais são utilizados pelas empresas e organismos mas, existem ainda muitos problemas relacionados com a sua modelação, como por exemplo, a incorreta utilização da notação ou modelação incorreta dos processos, muitas vezes podendo estas estarem relacionadas com a dificuldade de compreensão dos modelos.

A avaliação dos modelos será realizada através da recolha de métricas relativamente à complexidade e facilidade de compreensão dos modelos. Uma das maneiras de efetuar a recolha das métricas é através da extração de dados estatísticos dos modelos. Outra maneira, é a recolha métricas com recurso ao dispositivo *eye-tracker*, que permite recolher dados relativos a algumas reações faciais dos leitores dos modelos, tais como, o movimento dos olhos, a dilatação da pupila e a taxa de pestanejos do utilizador. Esta recolha de dados permite obter informações ao nível da compreensão dos modelos por parte das pessoas. A recolha destas métricas permitirá, perceber quais os elementos que provocam erros de interpretação ou maior dificuldade de compreensão [Sha+15].

Com o auxílio das regras de boas práticas de modelação definidas em [Cor14], segundo as normas em [OMG11], pretende-se ainda perceber se estas regras são de facto úteis e eficazes para a análise e compreensão dos modelos por parte dos seus utilizadores.

1.2 Objetivos do trabalho

A finalidade deste trabalho é desenvolver um *framework* que permitirá avaliar a complexidade e capacidade de compreensão dos modelos BPMN e qual o seu impacto na compreensão por parte dos utilizadores destes modelos. Foram definidas e implementadas um conjunto de métricas relacionadas com as informações retiradas diretamente dos modelos e métricas recolhidas com o auxílio de um dispositivo *eye-tracker*. Este por sua vez permitirá analisar o desempenho dos utilizadores na leitura e interpretação dos modelos.

Para recolher dados com o *eye-tracker* serão realizados testes com participantes, de modo a recolher dados reais da utilização e análise dos modelos por parte das pessoas. Por outro lado, as métricas serão implementadas em *Object Constraint Language* (OCL) e serão aplicadas em modelos reais, como os modelos de teste e eventualmente modelos fornecidos por alunos ou disponíveis em repositórios públicos.

As métricas implementadas no âmbito deste projeto serão validadas recorrendo a um conjunto de casos de estudo reais e a propriedades teóricas. De uma maneira mais detalhada, a definição e utilização de métricas tem o intuito de influenciar a criação de modelos, para que possuam uma melhor compreensão e, conseqüentemente, poderá encontrar defeitos da sua modelação, mitigando assim os erros e falhas cometidos durante a fase de modelação. Mas é preciso ter em atenção, que nem sempre a existência de dificuldades de compreensão de determinadas regiões dos modelos corresponde à existência de erros ou defeitos, a complexidade existente pode dever-se a uma grande complexidade essencial do processo.

1.3 Principais Contribuições Previstas

A principal contribuição desta dissertação consiste na definição e recolha de métricas para a avaliação dos modelos de processos de negócio **BPMN**, de modo a avaliar a complexidade do modelo e a sua facilidade de compreensão por parte dos seus utilizadores. Também se pretende avaliar os modelos de acordo com os seus diferentes utilizadores, ou seja, se estes são profissionais, ou caso contrário, se não possuem conhecimento dos modelos **BPMN**. Com esta avaliação, pretende-se influenciar uma melhor modelação, como por exemplo, evitar o uso dos elementos mais complexos e incertos quanto à sua interpretação.

A utilização do *eye-tracker* trás a grande vantagem de permitir uma recolha de métricas dinâmicas. A junção das métricas recolhidas dos modelos, pelo *eye-tracker*, com a recolha do tempo de realização das tarefas por parte do utilizador e, conseqüentemente, a correção das respostas dadas destes, fornecem informação que nos permitirá tentar responder às seguintes perguntas:

1. Qual o esforço de compreensão dos modelos **BPMN** consoante a sua complexidade?
2. Quais os elementos com maior complexidade de compreensão por parte dos utilizadores?

Pretende-se que estas sejam respondidas através de uma avaliação prática das métricas implementadas e, posteriormente, validadas. Antes de poder ser feita uma medição e validação, as métricas serão definidas segundo a abordagem *Goal Question Metric (GQM)* [Bas+94] e, de seguida serão definidas com recurso a dois géneros de linguagens: natural e formal. A formalização será conseguida utilizando a linguagem de programação **OCL** e esta formalização pretende eliminar as ambiguidades criadas e existentes em qualquer linguagem natural. Devido ao facto que já existem diversas métricas em **OCL** para os modelos **BPMN**, as métricas implementadas serão uma mistura de métricas novas com as já existentes.

1.4 Estrutura do Documento

Os próximos capítulos deste documento encontram-se estruturados da seguinte forma:

- **Capítulo 2 - Enquadramento** → Neste capítulo são apresentados e descritos os conceitos do contexto da dissertação, relacionados com a área de **ES, SI** e os *Business Process Modeling (BPM)*. Também é descrito o mecanismo de definição de métricas e, conseqüentemente, é apresentado o meta-modelo utilizado para a definição das mesmas. Ainda é detalhado o funcionamento do dispositivo *Eye-tracker*, como é realizada a recolha de dados e quais as suas limitações.

- **Capítulo 3 - Trabalho Relacionado** → Neste capítulo são apresentados alguns trabalhos relacionados com a avaliação dos modelos de processos de negócio e a recolha de dados utilizando o *Eye-tracker*.
- **Capítulo 4 - Métricas de avaliação dos Modelos BPMN** → Este capítulo é dedicado à apresentação da solução proposta para realizar uma avaliação da complexidade e compreensão dos modelos BPMN, para tal é apresentado um conjunto de métricas através da abordagem GQM.
- **Capítulo 5 - Implementação do Protótipo** → Este capítulo contém as possíveis soluções para a solução proposta e consequentemente qual a opção tomada e com base em que justificação. São também apresentadas as tecnologias utilizadas e um esquema da arquitetura da solução.
- **Capítulo 6 - Validação Teórica** → O objetivo deste capítulo é apresentar a validação teórica das métricas de complexidade implementadas para avaliar os modelos BPMN.
- **Capítulo 7 - Validação Empírica** → Neste capítulo são apresentados todos os detalhes e procedimentos para à validação empírica dos modelos e dos testes com o *eye-tracker*. São ainda apresentadas as hipóteses retiradas das perguntas colocadas na solução proposta.
- **Capítulo 8 - Análise dos Dados Recolhidos** → Este capítulo apresenta todos os dados recolhidos e as suas respetivas análises e comparações. É ainda neste capítulo que são dadas as respostas às hipóteses colocadas, de modo a retirar ilações dos dados e responder às perguntas inicialmente propostas.
- **Capítulo 9 - Conclusões** → É neste capítulo que são apresentadas as contribuições reais desta dissertação, assim como quais as limitações encontradas ao longo do trabalho. No fim são apresentados quais os trabalhos futuros a desenvolver, assim como quais aqueles que ainda estão em andamento.

ENQUADRAMENTO

Esta secção tem como objetivo introduzir os conceitos base utilizados para o desenvolvimento desta dissertação. Alguns destes conceitos têm por base a Engenharia de Software e os Sistemas de Informação. É ainda feita uma introdução aos modelos BPMN, ao nível da sua área de atuação e simbologia. Finalmente é descrito como é realizada a especificação formal das métricas.

2.1 Modelação de processos de negócio

Na modelação de processos de negócio, é importante perceber no que consiste o fluxo dos processos (do inglês *workflows*). Para qualquer tipo de desenvolvimento de *software* é necessário que tenha existido uma gestão/planeamento do fluxo dos processos do sistema [Wes10]. De uma maneira mais detalhada, é necessário que exista uma análise ou descrição das tarefas dos processos que devem ser modeladas. Também é bastante importante, perceber a troca e criação de informações por parte das tarefas [Som16].

Depois de especificados, podem então ser modelados segundo as normas dos modelos de processos de negócios, que por sua vez são modelos gráficos. Estes podem ser modelados usando: os **Diagramas de atividades** do grupo *Unified Modeling Language* (UML), **BPMN**, *Software Process Engineering Metamodel* (SPEM), *Business Process Execution Language* (BPEL). Para além da existência destes modelos, existem ainda empresas que criaram as suas próprias linguagens de modelação, a *OutSystems* possui a linguagem *Business Process Technology* (BPT) e a *Microsoft* a linguagem **Nintex Workflow**. É apresentado no **Apêndice A** um pequeno exemplo de modelação dos diferentes géneros de modelação dos BPM.

Todas estas linguagens de modelação contêm as mesmas características de representação das suas funcionalidades e requisitos, para tal são utilizadas atividades/tarefas e

fluxos de controlo entre as atividades. Estes representam ainda a ocorrência de eventos e fazem a distinção entre os participantes.

Segundo[DB07] os mecanismos de modelação dos processos de negócios são uma abordagem sistemática para gerir e melhorar o negócio de uma organização através de uma gestão ativa, que coordena todos os aspetos do desenvolvimento de *software*, tendo em vista a eficiência e eficácia dos objetivos definidos. A construção dos modelos BPM possuem um ciclo de vida, à semelhança dos *softwares*, e este é composto pelas seguintes fases:

- **Estratégia** → Criação das bases de todo o projeto, ou seja, é desenvolvida e pensada toda a estrutura do sistema.
- **Design** → Encontro entre as necessidades da organização com os requisitos dos utilizadores do sistema. É nesta fase que deve ser garantida a eficiência dos fluxos dos processos.
- **Implementação** → Transição dos requisitos e processos dos modelos para o software do sistema, minimizando sempre que possível as perdas de informação.
- **Controlo** → Verificação e atualização da eficiência e requisitos do sistema.

Na [Secção 2.2](#) são descritos em detalhe os modelos BPMN, que são o centro desta dissertação e na [Secção 2.2.1](#) são apresentados os restantes modelos anteriormente apresentados.

Os modelos de processos de negócio relacionam-se com a ES, em particular com a área da Engenharia de Requisitos, uma vez, que este género de modelação ajuda na fase de elicitação de requisitos e no desenvolvimento dos produtos, incluindo os SI. Uma vez desenvolvidos os modelos de processos de negócio farão parte do novo modelo de processos que será melhorado com os seus sistemas de informação e elicitação de requisitos.

2.2 BPMN

Os modelos BPMN trouxeram um grande avanço para a modelação dos processos de negócio, uma vez que esta é uma linguagem gráfica e independente, que tem a vantagem de poder ser usada de um modo eficiente tanto pela indústria empresarial como para a indústria das tecnologias de informação [Sha+11]. Estes juntam os processos de modelação orientados a objetos e o design da execução dos processos. De uma maneira mais específica, esta abordagem permite detalhar a análise de desempenho, especificação de requisitos e ainda o fluxo de execução das atividades e funcionalidades necessárias.

Este tipo de modelo é bastante rico em termos de simbologia, mas estes devem ser usados com moderação. Para apresentar uma boa modelação não é necessária a utilização de todos os símbolos disponíveis [OMG11; Whi08].

Podemos dividir os elementos de um modelo BPMN em cinco categorias:

- **Flow Objects** → Engloba os eventos, atividades e *gateways*.
- **Dados** → Composto por qualquer tipo de dados a representar tais como: objetos de dados, dados de entrada, dados de saída e armazenamento de dados.
- **Connecting Objects** → Composto pelos fluxos de sequência, fluxo de mensagens e associações.
- **Swimlanes** → Mecanismo de agrupamento dos elementos do modelo, podendo este ser através da utilização de *Pools* ou *Lanes*.
- **Artefactos** → Possuem a função de fornecer informação adicional para um processo, e recorre-se à utilização de grupos ou anotações de texto.

No [Anexo II](#) é possível encontrar uma documentação detalhada com toda a simbologia dos elementos destes modelos. Os elementos mais utilizados nos modelos [BPMN](#) são:

- **Atividades** → Trata-se da descrição de uma tarefa em específico. As tarefas não possuem quaisquer outros tipos de elementos internos, caso isso aconteça deixa de ser uma tarefa e passa a ser um processo.
- **Processos** → Como já referido, este elemento é a composição de diversas atividades. Estes podem pertencer a um dos seguintes subtipos de processos:
 - **Embutidos** → Usados para exemplificar de uma maneira clara um determinado fluxo, ou atividades.
 - **Reutilizáveis** → Invocados para dentro de um determinado modelo, e não possuem a necessidade de serem especificados neste. Tornam-se bastante úteis para a simplificação do modelo e abstração de determinados processos.
- **Pool e Lane** → Dentro de uma *Pool* podem existir as *Lanes* que forem necessárias, ou por outras palavras, podem existir os atores/intervenientes que forem necessários para a realização do processo.
- **Gateways** → Utilizados com o objetivo de unir e dividir o fluxo do modelo. Estes oferecem uma maior flexibilidade ao modelo. Os diferentes tipos de *gateways* são:
 - **Gateway Exclusivo** → O fluxo segue um único caminho.
 - **Gateway Inclusivo** → O fluxo segue pelo menos um caminho.
 - **Gateway Paralelo** → Obrigatoriedade de passar por todos os caminhos e cumprir os seus fluxos de execução.
 - **Gateway Complexo e Default** → Semelhante ao funcionamento de uma condição *if*, o fluxo do modelo segue o ramo em que a condição é verificada.

- **Eventos** → Expressão uma determinada ocorrência ou invocação no modelo. Existem três tipos de eventos, sendo que dentro de cada tipo de evento existem subtipos do mesmo.
 - **Eventos de início** → Declaração do início de um novo fluxo do modelo. Os subtipos destes eventos podem ser eventos de tempo, função ou mensagens.
 - **Eventos Intermédios** → Desencadeamento de uma ação, por outras palavras, é usado para representar a ativação de um novo fluxo no modelo ou quando alguma atividade invocou um determinado evento, que por sua vez terá as suas próprias atividades ou processos. Os subtipos também podem ser de tempo, função, mensagens ou ainda eventos ininterruptos.
 - **Eventos Finais** → Usados para por fim a um determinado fluxo existente no modelo. Os subtipos também podem ser de tempo, função, mensagens e ininterruptos.
- **Dados** → Usados para transmitir informações geradas ou existentes num elemento ou processo.
 - **Data Objects** → Mecanismo que permite a transmissão de dados entre diferentes atividades. A visibilidade deste género de dados depende de onde são criados.
 - **Coleção de Dados** → Transmissão de um conjunto de dados do mesmo tipo.
 - **Armazenamento de dados** → Dados persistentes no processo, ou seja, são visíveis durante toda a execução dos fluxos do processo.
- **Mensagens** → Representam a troca de informações entre diferentes processos do modelo, ou entre diferentes participantes do modelo.

Na [Figura 2.1](#) é possível visualizar um exemplo de um modelo BPMN. Este representa o registo de um novo membro na plataforma *Business Process Incubator (BPI)*. O processo começa com a finalização do preenchimento do formulário e a ativação da criação de uma nova conta. Independentemente da escolha por parte do utilizador sobre a subscrição de notícias, é pedido ao utilizador uma confirmação do email anteriormente fornecido. Após a confirmação do mesmo, o utilizador recebe uma notificação no seu email e pode então proceder ao *login* na plataforma.

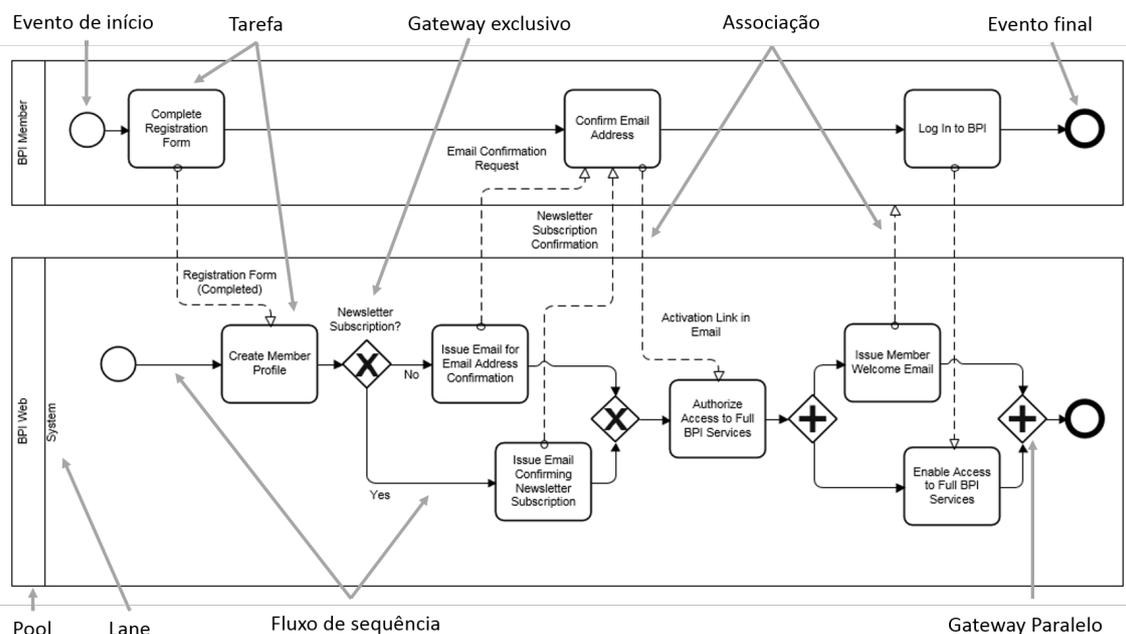


Figura 2.1: Exemplo de um modelo BPMN retirado de [Inc17]

2.2.1 Outras notações

De seguida, são apresentadas outras notações existentes para modelar os processos de negócio.

- **Diagramas de atividades** → Os diagramas de atividades **UML** são baseados e especificados através da passagem entre atividades. Estas recebem um ou mais fluxos de controlo e podem gerar um ou mais fluxos de controlo, podendo ou não modificar o estado do sistema. É possível modelar o envio e a receção de notificações externas nestes diagramas, assim como a possibilidade de usar temporizadores [DTH01; Rus+06]. Este género de diagramas são usados principalmente quando se pretende especificar e dar ênfase às diversas etapas que um processo possui [DTH01].
- **SPEM** → Os diagramas **SPEM** são modelos, onde a modelação foca-se na modularidade e flexibilidade em criar novos processos. Em termos de modelação estes também são orientados em atividades mas permitem a existência de diversas dependências entre os elementos dos processos [Ste+]. Este género de diagramas são bastante utilizados no que toca à modelação de processos de *software* [Ben+07].
- **BPEL** → O método de modelação **BPEL** é mais utilizado como mecanismo de modelação dos processos de negócios para *web services*, em que cada processo **BPEL** é acessível os seus utilizadores como um *web service*. Estes modelos são implementados recorrendo a linguagens baseadas em *eXtensible Markup Language (XML)* [And+03].
- **BPT** → A *outsystems* recorre aos modelos **BPT** para modelar os seus processos de negócios. Em particular esta linguagem permite um desenvolvimento modular, tanto

o *design* como a implementação das várias etapas dos processos estão incorporados nos modelos isto devido à plataforma *OutSystems Service Studio*, onde estes modelos são produzidos. A grande vantagem desta metodologia é o fabrico de diagramas de fácil perceção e reutilizáveis [Hen16].

- **Nintex Workflow** → À semelhança dos BPT estes diagramas da *Microsoft* são também concretizados numa plataforma própria, a *SharePoint*. Estes modelos têm como objetivo implementar e autonomizar os processos de negócios através de fluxos de controlo [Alb14; Cha+16].

Para todas as notações acima apresentadas, é apresentado no Apêndice A um exemplo com a sua respetiva explicação.

2.3 Sistema de Informação

Como já referido os Sistemas de Informação (SI) são desenvolvidos com as técnicas de desenvolvimento da ES e podem ser definidos como sendo uma combinação de pessoas, *hardware*, *software*, redes de comunicação, dados e políticas e procedimentos de armazenamento, acesso, transformação e disseminação de dados dentro de uma organização. Estes podem ser usados tanto para ambientes computacionais ou não computacionais [OM08].

As tecnologias de informação são uma área cada vez mais relevante para todos os setores económicos, sociais e governamentais nos dias de hoje [Bra05]. O grande desafio desta área é conseguir que o *software* desenvolvido contenha um bom desempenho no que toca às informações que devem ser geradas e guardadas. A qualidade dos sistemas de informações depende das técnicas utilizadas nas fases de desenvolvimento de *software*.

Os sistemas de informação são compostos por três componentes: dados, ações e utilizadores [Spr13]. Os *stakeholders* devem especificar o contexto da informação e quais as informações necessárias para o *software* em causa. Os dados existentes no sistema, devem ser relevantes, ou seja, se os dados determinam o fluxo das ações, estes devem ser guardados e atualizados. Consequentemente, todos os dados demonstrados pelo sistema serão visualizados e analisados pelo utilizador final e este, por sua vez, será um mecanismo de avaliação do sistema em si.

2.3.1 Engenharia de Software

A Engenharia de *software* (ES) é a área da engenharia que está encarregue de todos os aspetos ligados à produção de *software* desde as suas fases iniciais, ou seja, esta acompanha o desenvolvimento de um produto desde que é uma ideia até à sua concretização e, posteriormente, a sua manutenção durante o seu tempo de vida [Som16].

O processo de desenvolvimento de software é composto por quatro atividades fundamentais:

1. **Especificação de *software*** → Parte inicial de um projeto, onde os *stakeholders* e os engenheiros definem as funcionalidades e restrições do *software* em causa.
2. **Desenvolvimento de *software*** → Fase de implementação do software anteriormente planeado e especificado.
3. **Validação do *software*** → Verificação de que o software desenvolvido corresponde ao esperado, ou seja, se o resultado final apresenta todas as funcionalidades e requisitos especificados corretamente.
4. **Evolução do *software*** → Fase de manutenção e melhorias ao projeto após o seu desenvolvimento, com adição de novas funcionalidades e requisitos antes não previstos ou com a melhoria do desempenho do produto.

Apesar de todos os projetos conterem as atividades apresentadas, existem vários géneros de abordagens para o desenvolvimento de *softwares*, sendo que cada um apresenta o seu fluxo e etapas. Como existem diversos modelos, cada um adapta-se melhor a determinadas abordagens de desenvolvimento do *software*, sendo que por vezes, a melhor opção é a junção de diferentes modelações e abordagens para o desenvolvimento de *software*.

Um dos géneros de *softwares* descritos e desenvolvidos com as técnicas da **ES** são os **SI**, ou seja, recorre-se às metodologias de desenvolvimento da **ES**, para criar os sistemas de informação.

2.4 Controlo de Qualidade de software

Qualquer software que é desenvolvido ou modelado deve ser testado, através de mecanismos de avaliação. Estes testes e medições, ajudam a melhorar e avaliar a qualidade dos produtos de *software* desenvolvido. Para uma medição ser realista, deve centrar-se na especificação dos objetivos, ser aplicada em todas as fases de desenvolvimento e ser realizada segundo o contexto alvo do *software* [Bas+94].

2.4.1 Atributos de qualidade

Os atributos de qualidade relacionam-se com as propriedades do *software* e são aplicáveis a todos os sistemas computacionais e produtos de *software*. As propriedades são as seguintes apresentadas [Jun+04]:

- **Funcionamento** → Adequação do produto em relação às suas funcionalidades.
- **Fiabilidade** (do inglês *Reliability*) → Capacidade do produto corresponder ao que lhe é esperado durante um determinado período de tempo. Um sistema ou produto torna-se confiável se este for tolerante a falhas e estiver disponível ao utilizador.

- **Desempenho** → O desempenho relaciona-se com a quantidade de recursos utilizados. Este pode ser medido de acordo com o tempo de resposta e com a utilização dos seus recursos.
- **Usabilidade** → Esta característica consiste na capacidade do utilizador conseguir alcançar os seus objetivos de uma maneira eficaz, eficiente e com satisfação.
- **Segurança** → Consiste na proteção dos dados por parte do produto e como este lida com os seus diferentes níveis de autorização.
- **Compatibilidade** → Capacidade de troca das componentes de um produtos por outras, sem que as suas funcionalidades deixem de funcionar corretamente.
- **Capacidade de Manutenção** (do inglês *Maintainability*) → Grau de eficácia e eficiência durante a modificação do produto.
- **Portabilidade** → Eficiência e eficácia de transportar o produto de ambiente sem causar problemas ao seu normal funcionamento.

Apesar de todas as propriedades serem importantes, nesta dissertação, será dado uma maior ênfase à propriedade da usabilidade, ou seja, pretende-se analisar e avaliar a complexidade dos modelos e a capacidade de compreensão dos modelos por parte utilizadores, englobando assim a eficácia e eficiência tanto dos modelos como dos utilizadores.

2.4.2 GQM

A abordagem GQM [Bas+94] é uma metodologia de avaliação e medição de erros, defeitos, objetivos e desempenho de produtos da ES já existentes em ambientes particulares. Também pode ser utilizada para melhorar a qualidade dos produtos, definindo novos objetivos. Esta pode ser decomposta em três níveis de abstração:

- **Nível conceptual - Objetivo (GOAL)** → Definição dos objetivos a alcançar.
- **Nível operacional - Questão (QUESTION)** → Criação de um conjunto de perguntas que iram especificar os objetivos a alcançar.
- **Nível quantitativo - Métrica (METRIC)** → Definição de métricas a serem utilizadas para responder às questões definidas.

Para alcançar os objetivos definidos serão avaliados produtos, processos e recursos. Os produtos correspondem a artefactos ou documentos gerados durante o ciclo de vida do sistema. Os processos correspondem às atividades relacionadas com o sistema. E por fim os recursos são os *items* necessários para que os processos funcionem corretamente.

Este tipo de abordagem é hierárquica, como podemos observar na Figura 2.2. Para cada sistema, existirá a definição de vários objetivos, sendo que cada um deles terá um

conjunto de perguntas às quais se pretende responder. Cada questão será composta por um conjunto de métricas, podendo ser utilizadas para responder a mais que uma questão. Este tipo de abordagem é o recomendado pelo *Instituto de Engenheiros Eletricistas e Eletrônicos (IEEE) Computer Society* [IEE98].

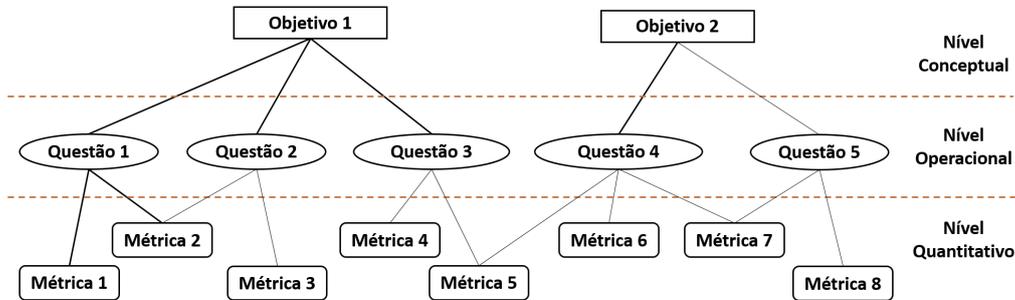


Figura 2.2: Modelo de abordagem GQM

2.5 Especificação de Métricas

A especificação das métricas pode ser realizada recorrendo a uma linguagem natural e a uma linguagem formal.

- **Linguagem natural** → Esta é uma técnica bastante usada, devido ao facto de ser intuitiva e universal, mas traz consigo alguns problemas, sendo estes a possibilidade da definição da métrica ser vaga ou ambígua [Som16].
- **Linguagem formal** → A definição de métricas através de uma linguagem formal evita ambiguidades, aumentando assim a sua consistência e permite ainda criar mecanismos de automatização na sua definição. Estas definições têm como base a utilização de conceitos matemáticos e lógicos. Por outro lado, uma definição formal exige conhecimento sobre a linguagem usada, para que seja compreendida [Abr01].

Para definir as métricas nesta dissertação, serão utilizadas ambas as linguagens acima apresentadas, sendo que para a definição formal será utilizada o OCL.

2.5.1 OCL

A linguagem OCL é utilizada para apresentar a descrição das métricas numa forma rigorosa. Esta trata-se de uma linguagem declarativa, sendo utilizada para representar e implementar expressões sobre os modelos UML [OMG14], nesta dissertação, serão criadas expressões sobre o meta-modelo do BPMN, que é definido através de um diagrama de classes.

Para além do OCL ser utilizado para especificar expressões, pode ser utilizada como uma linguagem de pesquisa, para especificar invariantes das classes, pré e pós condições

das operações e restrições. Também é possível descrever mensagens, ações e novas regras derivadas dos atributos dos modelos [OMG14].

Os grandes benefícios existentes com a utilização desta linguagem são: documentação melhor elaborada, uma vez que existe mais informação descrita; existência de uma maior precisão e sem ambiguidades, ou seja, a utilização da linguagem OCL, remove a divergência de interpretação da informação apresentada [Alm14] [OMG14].

Esta linguagem contém quatro tipos primitivos de dados: **Booleanos**, **Inteiros**, **Reais** e **Strings**. Estes tipos por sua vez, são interrogados ou analisados com recurso aos operadores lógicos existentes, que são: > (maior), < (menor), >= (maior ou igual), <= (menor ou igual) e = (igual). Para além destes operadores, existem ainda palavras-chave, que representam expressões condicionais sobre os atributos existentes. As palavras-chaves existentes são: *if-then-else*, *and*, *or*, *xor*, *not*, *implies*, *let-in* e *in*. Para além dos tipos de dados primitivos, existem ainda: **conjuntos**, podendo ser ordenados ou não, **sequências** e **bags** [OMG14].

No **Apêndice B** encontra-se um exemplo OCL de definições e restrições a partir de um diagrama de classes.

2.5.2 Complexidade

Existem diversas perspetivas complementares sobre a complexidade, algumas destas relacionam-se com os elementos do modelo em si [Car05b], ou seja, a complexidade corresponde à quantidade de elementos do modelo. Por outro lado, a complexidade pode relacionar-se com a dificuldade de compreensão da notação dos modelos e as suas técnicas de modelação[Edm95].

A complexidade de um diagrama cresce de acordo com o aumento de elementos representados no diagrama. Ao aumentar a complexidade pode-se aumentar também o esforço necessário para a compreensão do diagrama, assim como o esforço necessário para a manutenção e modificação do mesmo. [Car06; GL06b; Men+07; PC17].

Devido à complexidade crescente dos diagramas, esta passou a ser uma propriedade que deve ser medida e analisada. Para analisar a complexidade são utilizadas e criadas novas métricas, de modo a perceber e mitigar a sua complexidade. As métricas devem medir os diversos aspetos e componentes dos diagramas, assim como devem ser utilizados diferentes géneros de métricas [GL06b; SG+10].

Os aspetos avaliados pelas métricas podem ser agrupados em diferentes categorias [Car+06], sendo estas:

- **Atividades** → Quantidade de atividades presentes no diagrama.
- **Fluxo de controlo** → Quantidade de fluxos existentes no diagrama, entre as diversas atividades.
- **Fluxo de dados** → Relação entre os dados transferidos através do fluxo de controlo do diagrama.

- **Recursos** → Recursos externos necessários para o processo representado no diagrama ser concluído.
- **Profundidade** → Quantidade de decisões que devem ser tomadas no fluxo de controlo de modo a concluir uma determinada ação.
- **Saltos** → Fluxo de controlo de entrada ou saída para regiões diferentes dos modelos.
- **Peso cognitivo** → Soma do esforço necessário para compreender os elementos do diagrama, baseando estes valores em estudos empíricos (como por exemplo, experimentos ou inquéritos).
- **Anti-Padrões** → Soma das secções dos diagramas que não sigam os padrões, definidos por regras de boas práticas de modelação.

Dado o facto de os modelos **BPMN** serem gráficos, a capacidade de medição de uma métrica em processos interativos ou a sua modularidade também são aspetos importantes na análise da complexidade dos modelos [Edm95].

A validação de uma métrica pode ser decomposta em três passos [Muk10]: primeiro a métrica deve ser definida, de seguida esta passa por uma validação teórica e no final é feita uma validação empírica. Uma validação teórica baseia-se na verificação da estrutura da métrica, se esta é correta e adequada para o objetivo proposto. O mais importante nesta fase é perceber qual é o tipo de escala (como por exemplo, nominal, percentagem, intervalo de valores, entre outros) que deve ser utilizada para fazer a medição em causa. Por outro lado, a validação empírica verifica se as métricas permitiram alcançar o objetivo para o qual foram definidas. É nesta fase que são utilizados inquéritos ou são realizadas experiências de modo a fazer a validação da métrica [PC17].

Ainda em relação à validação teórica, são utilizados três métodos para a fazer, se a escala usada é a adequada, se as propriedades do *framework* de Briand são respeitadas e ainda se as propriedades de Weyuker também são respeitadas. Estas propriedades estão detalhadas na [Secção 2.5.3](#).

Na literatura são apresentadas sessenta e seis métricas [PC17], retiradas de diversos artigos, apresentadas no [Anexo III](#). Quanto às métricas apresentadas, o domínio dos processos de negócio é aquele que possui mais diversidade de métricas, sendo estas na sua maioria relacionadas com a notação gráfica utilizada. A maioria das métricas avaliam um ou dois tipos de componentes (atividades, fluxo de controlo, fluxo dos dados ou recursos). A componente que é mais avaliada é o fluxo de controlo, com vinte e seis métricas diferentes e o par de componentes mais avaliado é a junção das atividades com o fluxo de controlo. Isto acontece devido ao facto de serem as componentes mais utilizadas nos diagramas.

2.5.3 Propriedades e categorização das métricas

Weyuker propõe nove propriedades que oferecem um ferramenta de avaliação para as métricas definidas. As propriedades são [Bri+96; Car05a]:

1. Medições de processos diferentes não possuem o mesmo grau de dificuldade, ou seja, não é obtido o mesmo resultado.
2. A métrica deve conseguir demonstrar resultados, por outras palavras, a granularidade do que está a ser medido deve ser informativa de modo a retirar conclusões.
3. Para processos semelhantes quando aplicada a métrica, devem ser obtidos resultados semelhantes.
4. Mesmo que dois processos tenham a mesma funcionalidade, não implica que possuam a mesma complexidade.
5. A junção das complexidades de dois processos diferentes é sempre maior ou igual à complexidade do processo menos complexo.
6. Quando agregados dois processos será obtido um valor de complexidade, que por sua vez, será diferente do valor obtido da agregação de um destes processos com qualquer outro processo.
7. Mesmo que um processo seja a permutação da ordem das atividades de outro processo, não implica que o resultado das métricas quando aplicado seja o mesmo.
8. A mudança de nome das atividades não interfere com os resultados obtidos das métricas.
9. Uma métrica deve ser capaz de detetar quando existem mudanças de comportamento no processo que é avaliado. Como por exemplo, a junção de vários processos pode resultar na obtenção de valores maiores na aplicação da métrica do que uma simples somas das complexidades dos mesmos processos.

Por outro lado, Briand categoriza as métricas em cinco grupos distintos [Bri+96; Muk10]. Para qualquer um destes grupos, os resultados obtidos nunca podem ser negativos e só é obtido um valor *null* quando não existem elementos no modelo. As categorias são:

1. **Tamanho** → Quantidade de elementos presentes no modelo, possuindo propriedade de adição.
2. **Comprimento** → Pode ser descrita como a distância entre dois elementos. A propriedade de adição não afeta os processos em si, mas sim o resultado obtido pela métrica.

3. **Complexidade** → Esta não pode depender da convenção usada para a representação da conexão entre os elementos. Se os processos forem disjuntos a complexidade é a soma da complexidade de todos os processos. Por último, caso exista interseção entre os processos, o resultado final da complexidade nunca será menor que a soma das complexidades dos processos.
4. **Acoplamento** (do inglês *coupling*) → Esta captura a relação entre os elementos de diferentes módulos.
5. **Coesão** → Utilizadas quando é necessário referir vários módulos e existe a necessidade de encapsulamento de alguns processos ou funcionalidades. Não existe propriedade de adição, ou seja, o resultado da junção dos módulos é o maior valor de coesão existente.

2.6 Eye-tracking

O *Eye-tracking* é um mecanismo de recolha de dados cognitivos dos seus utilizadores. Este mecanismo é utilizado para realizar estudos empíricos e para estudar os modelos de compreensão [Sha+15], para perceber o que pode ser melhorado para facilitar a interação dos produtos/sistemas para os seus utilizadores.

Sendo esta uma ferramenta tecnológica, podem ser feitas distinções de acordo com os modelos existentes. Ou seja, um *Eye-tracker* pode ser do tipo intrusivo ou não intrusivo. Os modelos intrusivos, apresentados na Figura 2.3, retiram as suas medições dos movimentos dos olhos e da cabeça, através das várias câmaras colocadas ao nível dos olhos do utilizador. A desvantagem destes modelos é a possível influência nos valores obtidos devido ao desconforto causado ao utilizador [Sha+15]. Sendo o modelo não intrusivo o escolhido para a recolha de dados nesta dissertação.

Os modelos não intrusivos, Figura 2.4, baseiam-se na recolha de dados da análise das imagens recolhidas com recursos às câmaras do dispositivo. Os dispositivos mais recentes, recorrem à posição do nariz, olhos, sobrancelhas e lábios para perceber o movimento os olhos [Sha+15].



Figura 2.3: Exemplo de um modelo *eye-tracker* intrusivo



Figura 2.4: Exemplo de um modelo *eye-tracker* não intrusivo

2.6.1 Recolha e Análise de Dados

Como anteriormente dito, o *eye-tracker* é utilizado para recolher dados sobre a capacidade de compressão do seu utilizador e para tal ser possível analisam-se os movimentos dos olhos, sendo que existem dois tipos de movimentos [Zag+16]:

- **Fixação** → Este consiste no foco de atenção num único ponto durante, pelo menos, duzentos a trezentos milissegundos. Um foco de atenção pode indicar a existência de uma área de interesse (do inglês *Area of Interest (AOI)*), relacionando estes focos com o processo cognitivo, estes podem indicar um aumento da atenção necessária para interpretar, memorizar ou compreender a tarefa em causa.
- **Sacadas** → Trata-se na troca do foco de atenção de duas localizações distintas, por norma, este movimento é realizado entre trinta a oitenta milissegundos. Com este tipo de movimentos, é possível medir a sua velocidade e comprimento. É ainda possível captar padrões destes movimentos, sendo estes chamados de *scanpaths*, que por sua vez são um tipo específico de Sacadas.

Para além destes dois movimentos, também é possível recolher informação, de acordo com [Zag+16]:

- **Dilatação da pupila** → Esta é uma reação involuntária por parte do nosso corpo, sendo que a dilatação da pupila pode ir de um milímetro e meio até oito milímetros de diâmetro. A dilatação da pupila relaciona-se com o esforço cognitivo, uma vez que quanto maior o esforço para realizar uma tarefa, maior é a sua dilatação. Nas experiências em que esta informação é tida em conta é preciso ter em atenção a luminosidade onde são realizados os teste de modo a influenciar o menos possível os dados recolhidos.
- **Taxa de pestanejos** (do inglês *Blink Rate*) → Pode ser um movimento voluntário, mas na sua grande maioria trata-se de uma movimento involuntário. Este tipo de medição é relevante quando se pretende medir a atenção do utilizador em relação a alguma tarefa, ou seja, quando existe uma pequena taxa de pestanejos, é possível inferir que o utilizador está mais focado e atento ao que está a realizar.

Com estas quatro propriedades é então possível perceber quais as partes dos modelos que provocam um maior ou menor esforço de compreensão pela parte do utilizador. De modo a completar a informação à cerca da recolha de dados, no [Anexo IV](#) são apresentadas as diversas métricas presentes em [Sha+15].

2.6.2 Limitações

A recolha de dados com um dispositivo *eye-tracker* sofre de algumas limitações, tais como [Sha+15]:

- **Precisão** → Relacionado com o próprio erro que o *eye-tracker* introduz, com o tipo de calibração feita e ainda com diversidade de feições que os participantes podem apresentar.
- **Efeito Hawthorne** → Provocado pela introdução de influências nos resultados por parte do utilizador. Pode acontecer durante os testes, ao investigador interagir com o participante dos testes, ao calibrar o dispositivo e finalmente ao não parar a recolha de dados entre os diferentes testes.

Para além destes, também deve haver o cuidado de minimizar a interferência de fatores estranhos como a utilização de ecrãs distintos [Sha+15], ou condições ambientais diversas (como por exemplo, a variação da luz ambiente) para minimizar os potenciais impactos nos dados recolhidos e posteriormente analisados.

2.7 Sumário

Neste capítulo, foram apresentados conceitos relativos à Engenharia de *Software*, Sistemas de Informações e Modelação de Processos de Negócio. Foram apresentadas algumas abordagens diferentes de modelação dos processos de negócio, dando ênfase aos modelos **BPMN**, por ser o foco desta dissertação.

Foram apresentados os métodos de controlo da qualidade destes modelos, ou seja, foi abordada a técnica **GQM** e para a formalização de métricas o **OCL**. Por fim, é apresentado o dispositivo *eye-tracker*, juntamente com as informações por ele recolhidas, que permitirão retirar ilações sobre a compreensão dos modelos por parte das pessoas.

TRABALHO RELACIONADO

Este capítulo apresenta trabalhos que pretendem avaliar a qualidade dos processos de negócio. Por fim, é apresentado um pequeno balanço de acordo com os trabalhos analisados.

3.1 Análise da qualidade dos modelos de processos de negócio

Existem alguns trabalhos cujo objetivo é analisar a qualidade e/ou complexidade dos modelos BPMN [Aya12; Bar+15; Cor14; FL11; GL06b; Khl+09; Mak+10; SG+12b]. De um modo mais detalhado, o objetivo central de todos os trabalhos analisados consiste em perceber as falhas e erros existentes na modelação dos modelos BPMN e nas suas ferramentas, para que futuras modelações possuam menos falhas e uma menor complexidade. Alguns trabalhos relacionam-se com a implementação de regras de modelação [Bar+15; Cor14]. As regras são decompostas em dois conjuntos complementares: as regras de boa modelação modelos e as de boas práticas de modelação. Mas na sua grande maioria, os trabalhos estão relacionados com a implementação de métricas relacionadas com a complexidade dos modelos [Aya12; Bar+15; FL11; GL06b; Khl+09; Mak+10; SG+12b].

Para proceder à especificação dos objetivos e à definição de quais os métodos a seguir para os alcançar os seus objetivos, alguns trabalhos recorrem à abordagem GQM [Aya12; Cor14; SG+12b], tal como a presente dissertação.

Para que as métricas possam ser recolhidas, por norma é definido um meta-modelo. O meta-modelo representa as informações relativamente aos elementos existentes na linguagem e as suas conexões, por exemplo: diagrama de classes UML. Relativamente aos meta-modelos definidos, estes podem [Bar+15; Cor14] ou não [Aya12; Mak+10] seguir a documentação oficial da *Object Management Group* (OMG) [OMG11]. Com o conjunto de métricas a recolher e o meta-modelo, pode-se então proceder à especificação das métricas,

que na sua grande maioria estão implementadas recorrendo ao OCL [Aya12; Bar+15; Cor14].

Em especial nos trabalhos relacionados com a implementação [Bar+15; Cor14] e criação [Cor14] de métricas das regras de modelação, é feita uma pesquisa das ferramentas existentes para definir os modelos. Esta pesquisa tem por objetivo perceber a capacidade das ferramentas validarem as regras de modelação [Bar+15; Cor14] e, conseqüentemente, retirar conclusões sobre as regras existentes na documentação oficial desta linguagem de modelação, ao nível da sua perceção, cumprimento e o que pode ser melhorado [Cor14]. Este género de análise torna-se interessante devido à existência de definições ambíguas na semântica oficial dos modelos, levando à implementação de ferramentas de modelação incompletas e/ou incorretas a nível semântico.

Apesar destes modelos serem bastante utilizados, relevam-se ainda algumas limitações, sendo que estas refletem-se na qualidade dos modelos. Quando introduzidas as regras de boa formalização dos modelos, estes sofrem um grande impacto pela positiva em relação à sua qualidade [Cor14]. Em grande parte, as ferramentas existentes falham na validação dos eventos e *gateways* em ambos os géneros das regras de modelação, mas são as regras de boa modelação que mais falham, ou seja, poucas ferramentas validam este género de regras [Bar+15].

Com a implementação das regras de modelação dos modelos BPMN é então colocada a pergunta [Bar+15]: O número de regras violadas está relacionado com o nível de complexidade do modelo? Sim, a quantidade de defeitos nas regras de boa formalização dos modelos influencia a sua complexidade, mas esta também está correlacionada com o tamanho dos modelos. Esta conclusão foi obtida com recurso à implementação de métricas e, posteriormente, à análise da variância ANOVA, avaliando o efeito das variáveis dependentes (número e diversidade de regras de boa formalização violadas, número e diversidade das regras de boas práticas violadas) e independentes (*ranking* de complexidade, tamanho, volume, dificuldade e modularização).

Um outro trabalho [Aya12], avalia a qualidade dos modelos segundo três géneros de erros que podem ocorrer num modelo:

- **Ambiguidade** → A ambiguidade ocorre quando existem elementos com o mesmo significado e uma maneira de ser evitada passa pela inserção de identificadores únicos.
- **Abstração** → A abstração consiste na existência de regiões dos modelos que não possuem a granularidade mais indicada e, como tal, devem ser substituídas ou melhoradas.
- **Imprecisão** → A imprecisão consiste na existência de modelações incompletas que pode ser combatida com o preenchimento das regiões em falta.

Estes três erros foram avaliados recorrendo à comparação de dois meta-modelos e especificação de métricas. O primeiro meta-modelo especifica a linguagem dos BPMN e o

segundo representa as ontologias destes, ou seja, as classes e as suas relações e, por fim, as interações entre os elementos e as reações geradas por estas interações.

Existem também trabalhos relacionados com a adaptação de métricas da área da Engenharia de *software* [GL06b; Mak+10] e dos conceitos da orientação a objetos [Khl+09].

As métricas de Engenharia de *software* podem ser divididas em três categorias diferentes: **Acoplamentos**, **Coesão** e **Complexidade** (podendo ser consultadas no Anexo III). O aspeto mais relevante neste trabalho [Mak+10] está na especificação do meta-modelo, sendo que este tem em conta quatro perspetivas:

- **Funcional** → A perspetiva funcional tem em conta quais os elementos dos processos que são utilizados e os fluxos correspondentes. Enquanto que a perspetiva Comportamental tem em conta quando é que os elementos são utilizados e como se relacionam, como por exemplo, como as decisões e ciclos são executados.
- **Comportamental** → A perspetiva comportamental tem em conta quando é que os elementos são utilizados e como se relacionam, como por exemplo, como as decisões e ciclos são executados.
- **Organizacional** → O ponto de vista organizacional preocupa-se onde e com que elementos executam os processos.
- **Informativa** → A perspetiva informativa relaciona-se com todas as entidades de informação geradas e manipuladas pelo processo.

Apesar deste trabalho implementar diversas métricas de avaliação e apresentar os seus valores consoante o modelo testado, estas não são implementadas com o objetivo de perceber ou retirar ilações dos modelos, o que não é o caso da presente dissertação, que pretende com o auxílio de métricas perceber os níveis de compreensão e complexidade dos modelos.

Já num outro documento [GL06b], são definidas métricas de complexidade e chegou-se à conclusão que os modelos devem ser bem estruturados a nível do encadeamento das atividades (medido com *Maximum nesting depth*, *Mean nesting depth* e *knot count*, que têm em conta os saltos ou junções do fluxo de controlo do modelo). Por outro lado, quanto maiores os valores do *fan-in* e *fan-out*, maior pode ser a complexidade e o esforço de compreensão para interpretar o modelo. É preciso ter em conta que este trabalho analisa unicamente os fluxos de controlo dos modelos.

As métricas de orientação a objetos são traduzidas para os BPMN através de uma correspondência quase direta [Khl+09]. Estas métricas tem a vantagem de retirar informações dos modelos permitindo avaliar os acoplamentos; coesão dos processos e das atividades. As métricas definidas estão divididas em dois grupos:

- **Métricas de Acoplamento** → No âmbito dos modelos **BPMN**, trata-se da relação ou conexão entre atividades. As métricas adaptadas são: **ICP** (Acoplamentos importadas de um processo), em suma é o número de mensagens enviadas pelo sub-processos em causa; **ECP** (Acoplamentos exportadas de um processo), contagem das mensagens recebidas pelo sub-processo; **RFP** (Acoplamentos do fluxo de controlo); **LDA** (Localização dos dados das atividades), agregação da quantidade de dados gerados por uma atividade e os dados utilizados por si, mas gerados por outras atividades;
- **Métricas de coesão** → Concentra-se na coesão entre as atividades e as suas informações de *input* e *output*. As métricas adaptadas são: **TPC** (Coesão *tight* do processo), percentagem de pares de tarefas diretamente relacionadas; **LPC** (Coesão *loose* do processo), percentagem de pares de tarefas diretamente e indiretamente relacionadas.

Um outro objetivo de avaliação dos modelos **BPMN** [FL11] é analisar a capacidade cognitiva recorrendo às diferentes relações existentes dos elementos, tais como: **Relações entre elementos**, **Interatividade dos elementos** e **Separação dos elementos**. Os elementos de um modelo relacionam-se com funcionalidade dos diferentes elementos, mais concretamente com: Ordenação das operações, Concorrência, Repetições e Exclusividade. A interatividade de um elemento é algo importante pois avalia a quantidade de elementos a que está ligado. Quanto maior a interatividade for, maior é o esforço cognitivo necessário para o perceber. Um caso especial da interatividade dos elementos, é a independência entre os elementos, ou por outras palavras, é quando a ligação dos elementos é única e caso seja removida, os elementos deixam de estar conectados.

Após a análise dos dados [FL11], também com recurso aos conceitos de variáveis dependentes e independentes e a análise de concorrência **ANCOVA**, chegou-se à conclusão que a relação entre os elementos influencia o grau de dificuldade cognitiva da perceção do modelo, mais concretamente, foi verificada que a relação com menor subjetividade foi a ordenação dos elementos, seguida pela concorrência. Por outro lado, a relação com menos respostas corretas foi exclusivamente sobre os elementos e, por fim, a repetição ou ciclos foi a mais difícil de acordo com os participantes. Ainda foi possível concluir que a distância entre os elementos influencia a dificuldade cognitiva da compreensão do modelo, ou seja, quantos menos fluxos de controlo existirem entre as diferentes atividades melhor é a perceção da relação entre elas.

Um outro mecanismo de avaliação é a definição de intervalos de valores (do inglês *thresholds*) [SG+12b], com recurso também à análise da variância e distribuições normais. Esta definição tem por objetivo extrair informação dos elementos dos modelos **BPMN**, ao nível da sua compreensão e capacidade de modificação. A não compreensão de um modelo, afeta diretamente a capacidade de modificação de um modelo. Um dos grandes fatores de complexidade de um modelo são relacionados com os *gateways*, também já

provado em [Bar+15]. Devido à complexidade destes elementos, a avaliação deste trabalho incidiu sobre estes e, para proceder à análise da complexidade destes modelos foram consideradas as seguintes medidas: **CFC** (Complexidade do fluxo de controlo), **GM** (Incompatibilidade de *gateways*), **GH** (Heterogeneidade dos *gateways*), **AGD** (Média do grau dos *gateways*), **MGD** (Grau máximo dos *gateways*), **TNG** (Número total de *gateways*). Para que possam ser retiradas ilações é ainda definida uma variável **GCI** (Indicador da complexidade *gateway*). Para a validação destes, foram utilizadas técnicas de *cross-validation*, tais como: cálculo da precisão, memória (do inglês *recall*), rigor (do inglês *accuracy*) e **ROC** (representação gráfica das características de receção dos dados, permitindo comparar e analisar o desempenho dos participantes).

Com os resultados obtidos [SG+12b], conclui-se que quanto maior for o valor obtido, menor é a eficiência dos modelos, ou seja, mais difícil se torna a compreensão e modificação dos modelos. Ainda é possível afirmar que o mecanismo de definição de intervalos é bom para a classificação dos modelos, uma vez que os valores de precisão, memória e rigor são elevados. Apesar deste ser um bom mecanismo de classificação, quando os modelos se tornam mais complexos, deixa de ser possível distinguir se os dados estão relacionados com a compreensão ou a modificação, mas estes não são os únicos casos em ter em conta na avaliação dos modelos.

Independentemente do mecanismo utilizado para proceder à recolha da dados (métricas ou intervalos de valores), devem passar por uma fase de validação teórica e empírica. Para a validação teórica recorre-se às propriedades de *Weyuker*, *Briand* ou a algum outro mecanismo de validação definido [Bar+15; Cor14], tendo sempre como base as características dos elementos dos modelos. A validação empírica é realizada com recurso ao *framework* que é definido, ou seja, esta é realizada com recurso aos dados obtidos através de experimentos [Aya12; Bar+15; Cor14; FL11; GL06b; Mak+10; SG+12b]. Os experimentos podem ser feitos com recurso a modelos disponíveis em repositórios públicos, modelados por estudantes [Aya12; Bar+15; Cor14; Mak+10] ou a modelos definidos pelos autores [GL06b]. Por outro lado, os experimentos podem ser feitos com recursos a voluntários e questionários [FL11; SG+12b].

Em contraste com a grande maioria dos trabalhos analisados, esta dissertação irá realizar uma validação teórica, seguindo ambas as propriedades apresentadas. Será realizado também os dois géneros de validações empíricas anteriormente apresentadas.

3.2 Análise cognitiva recorrendo ao Eye-tracking

Não foi encontrado nenhum trabalho de avaliação dos modelos **BPMN** especificamente, mas foram encontrados trabalhos relacionados com o seu mecanismo de recolha de dados [PM13; Pin+12] e com o processo de modelação dos modelos de processos de negócio [Pin+12].

A compreensão de um utilizador é bastante volátil [PM13], uma vez que depende da familiarização com a linguagem de modelação, com as suas metodologias de leitura ou

com as características dos modelos (tamanho, complexidade, notação). Em particular, este trabalho analisa a compreensão do modelo com base nas suas **Regiões Relevantes (RR)**, que serão extraídas com recurso ao *eye-tracker*. Algumas das métricas utilizadas para os testes relacionados com as regiões relevantes foram as seguintes: **Número de RR**, *Scan Path Precision (SPP)*, *Scan Path Recall (SPR)* e *Scan Path f-measure (SPF)*. Para um utilizador conseguir retirar conclusões sobre o modelo é necessário que este conheça, previamente a sua notação, tanto a nível semântico como sintático.

O procedimento de recolha de dados com o dispositivo *eye-tracker* também será utilizado na corrente dissertação, sendo que nesta o foco principal será perceber as regiões com mais dificuldade de compreensão. Após a análise de dados, observou-se que os elementos com uma maior taxa de fixação, tanto a nível de tempo, como da quantidade de utilizadores, está correlacionada com a probabilidade de fornecer uma resposta correta ao teste realizado. No que toca a regressão linear, é possível concluir que de facto existe uma correlação positiva entre o tempo de fixação nas regiões relevantes e a resposta fornecida.

Independentemente da fase do desenvolvimento de um modelo, quando o modelador enfrenta um problema, o primeiro passo consiste na representação do problema, para que este seja capaz de criar uma solução [Pin+12]. Uma vez que os problemas podem-se tornar complexos facilmente e que a capacidade do ser humano é limitada, o problema é dividido em vários outros problemas de menor dimensão. As medições utilizadas para medir a complexidade são: **Número de fixações** e **Duração média das fixações**. Um dos valores mais relevantes para a compreensão do processo a modelar é o tempo decorrido entre o momento de conhecimento do problema e o início da modelação. Para o âmbito deste trabalho, um dos tempos relevantes também poderá ser o tempo que um participante demora a dar uma resposta ao teste que lhe é colocado.

Após a recolha de dados do processo de modelação dos modelos [Pin+12], pode-se afirmar que em média, o utilizador passa mais tempo fixado no modelo do que na descrição de um problema. Outra conclusão é que, após o início do desenvolvimento da solução para o problema em causa, existem poucos intervalos, ou seja, o participante só inicia a modelação após ter uma ideia formada sobre o que deve realizar. Em relação à avaliação dos modelos através de questionários com voluntários, é provável que também se possa chegar a estas duas conclusões, ou seja, o utilizador passará mais tempo a interpretar o modelo e as perguntas colocadas.

3.3 Sumário

Neste capítulo, são descritos diversos trabalhos dentro do contexto desta dissertação. Estes relacionam-se com a avaliação da qualidade dos modelos dos processos de negócios e com a recolha de dados através do dispositivo *eye-tracker*.

MÉTRICAS DE AVALIAÇÃO DOS MODELOS BPMN

Neste capítulo são apresentadas as métricas implementadas para avaliar a complexidade de compreensibilidade dos modelos **BPMN**. Tal como já foi descrito ao longo do documento, a definição de métricas foi realizada seguindo a abordagem **GQM**.

4.1 Diagrama GQM

De modo a facilitar a compreensão da aplicação de uma abordagem **GQM**, para a definição de métricas, é apresentado um esquema na **Figura 4.1**, com as especificação dos diferentes níveis existentes na abordagem. De uma maneira sucinta, são apresentadas onze questões, com um total de sessenta e cinco métricas, sendo que destas tinta e cinco são novas (representadas a **laranja**) enquanto que as restantes trinta foram previamente definidas por outros autores.

O nível conceptual corresponde ao alvo desta dissertação, ou seja, corresponde à avaliação da complexidade dos modelos **BPMN**. Para que seja possível retirar ilações dos modelos sobre a sua complexidade, foram analisados dez casos concretos de modelação e um caso de memorização, representados no nível operacional do diagrama. De modo a que seja possível responder às perguntas colocadas, foram implementadas métricas, sendo estas uma mistura de novas métricas, com métricas já existentes. Por suas vez, as métricas encontram-se representadas no nível quantitativo do diagrama, as métricas **M** são obtidas unicamente dos dados dos modelos, enquanto que as métricas **ME** são obtidas recorrendo a um dispositivo *eye-tracker*.

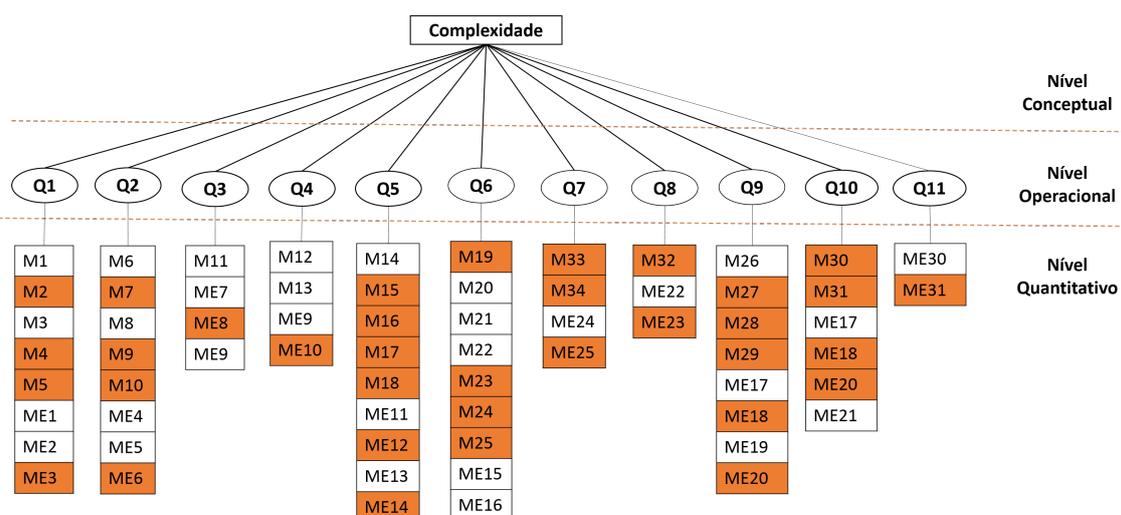


Figura 4.1: Modelo de abordagem GQM proposto

4.1.1 Objetivo - Nível Conceptual

O objetivo e nível conceptual da abordagem **GQM**, sendo que este é a avaliação da complexidade dos modelos **BPMN** e da sua facilidade de compressão.

4.1.2 Questões - Nível Operacional

Como anteriormente referido foram definidas onze perguntas, representativas do nível operacional da abordagem **GQM**. Estas são definidas para que seja possível conseguir recolher as informações necessárias para compreender o nível de compreensão dos modelos, sendo estas:

- **Q1** - De que modo a dificuldade de compreensão dos modelos aumenta com a quantidade de diferentes *gateways* utilizados?
- **Q2** - De que modo a dificuldade de compreensão dos modelos aumenta com a quantidade de diferentes eventos utilizados?

As questões **Q1** e **Q2** pretendem analisar se a quantidade e diversidade dos eventos e dos *gateways* existentes no modelo dificulta a compreensão destes, ou por outro lado, se os significados dos elementos em causa são perceptíveis através da visualização dos seus símbolos e nome.

É preciso ter em conta que com base nos trabalhos analisados, a questão **Q1** foi analisada anteriormente. No entanto, foi realizada com uma definição de intervalos de valores [SG+12b], não tendo interação entre leitores e modelos.

- **Q3** - De que modo a dificuldade de compreensão dos modelos aumenta com o aumento da interação entre participantes do processo?

- **Q4** - De que modo a dificuldade de compreensão dos modelos aumenta com a quantidade de participantes no processo?

As questões **Q3** e **Q4** pretende aferir a compreensão dos modelos baseando-se nos participantes do processo que é modelado, ou seja, se é perceptível a mudança de participante no decorrer do processo e a respetiva interação entre os diversos participantes.

- **Q5** - De que modo a dificuldade de compreensão dos modelos aumenta com o aumento da granularidade dos dados transmitidos entre as atividades do processo?

A questão **Q5** tenta relacionar a perceção da transmissão de dados entre participantes, atividades ou sub-processos e consequentemente se a as diferentes granularidades afetam a compreensão dos dados que são transmitidos.

- **Q6** - De que modo a dificuldade de compreensão dos modelos aumenta com o desequilíbrio entre a quantidade de fluxos de controlo dos elementos do processo?
- **Q7** - De que modo a dificuldade de compreensão dos modelos aumenta com a existência de fluxos paralelos de participantes no processo?
- **Q8** - De que modo a dificuldade de compreensão dos modelos aumenta com a existência de ciclos no processo?

As questões **Q6**, **Q7** e **Q8** pretendem analisar os fluxos de controlo de um processo de três modos diferentes. Em primeiro lugar pretende-se analisar o *fan-in* e o *fan-out* dos diferentes elementos no modelo. Em segundo lugar se a subdivisão do fluxo de controlo influencia a capacidade de análise e compreensão do processo a ser modelo. Por último, pretende-se perceber se a repetição de determinadas atividades, presentes nos ciclos, são claras e fáceis de compreender por parte do utilizador.

As questões **Q6** e **Q7** não foram colocadas em trabalhos anteriores, mas foram implementadas métricas que podem ser usadas para responder a estas questões. O trabalho que implementa as métricas [Khl+09], analisa modelos sempre com a mesma complexidade, não obtendo resultados quantitativos da complexidade dos elementos avaliados.

- **Q9** - De que modo a dificuldade de compreensão dos modelos aumenta com o desequilíbrio entre o número de elementos dos sub-processos ?
- **Q10** - De que modo a dificuldade de compreensão dos modelos aumenta com a existência de decomposição do processo em vários sub-processos?

As questões **Q9** e **Q10** analisam a compreensão dos sub-processos por parte dos leitores dos processos, ou seja, se o leitor sabe identificar quando existem sub-processos e se a decomposição de um processo em vários sub-processos afeta a compreensão do processo por parte dos seus leitores.

À semelhança da questão Q6 e Q7 a questão Q8 também não foi colocada em trabalhos anteriormente desenvolvidos, mas foram implementadas métricas que podem ser usadas para responder a estas questões [Khl+09].

- Q11 - Os significados e funcionalidades dos elementos dos modelos são fáceis de perceber e recordar?

A questão Q11 pretende analisar os modelos a nível da sua modelação e linguagem, por outras palavras, se a maneira de modelar afeta a compreensão dos modelos e se a vasta linguagem dos modelos é intuitiva e fácil de memorizar por parte dos seus utilizadores.

4.1.3 Métricas - Nível Quantitativo

Para cada uma das perguntas apresentadas é definido um conjunto de métricas que permitirá alcançar os objetivos da avaliação da complexidade e da avaliação da compreensão dos modelos. Estas métricas representam do nível quantitativo da abordagem GQM.

A Tabela 4.1 representa as métricas correspondentes à pergunta em causa, sendo que nesta tabela, a primeira coluna corresponde às questões, enquanto que a segunda coluna corresponde às métricas definidas para responder às questões. De um modo mais detalhado, a segunda coluna apresenta as métricas relacionadas unicamente com os modelos BPMN (M) e as métricas a serem recolhidas com o auxílio de um dispositivo *eye-tacker* (ME) e, para que se possa recolher dados foram realizados testes de interpretação e leitura dos modelos por parte de voluntários.

No que diz respeito à análise e recolha do tempo de fixação, também foi recolhida informação referente à dilatação da pupila e à percentagem de pestanejos por parte do utilizador, durante o decorrer do tempo de fixação. As métricas da questão Q11 são retiradas de [Cor14] e são apresentadas no Anexo V.

Tabela 4.1: Conjunto de métricas para a avaliação da compreensão dos modelos BPMN

Nível Quantitativo - Métricas	
Q1	M1/M6 - Número de <i>gateways</i> /eventos
Q2	M2/M7 - Percentagem de <i>gateways</i> / eventos
	M3/M8 - Número de elementos de cada género de <i>gateways</i> /eventos
	M4/M9 - Percentagem de elementos de cada género de <i>gateways</i> /eventos
	M5/M10 - Número <i>gateways</i> /eventos distintos
	ME1/ME4 - Tempo de fixação por tipo de <i>gateways</i> /eventos
	ME2/ME5 - Número de movimentos sacadas para <i>gateways</i> /eventos
	ME3/ME6 - Percentagem do tempo de fixação em <i>gateways</i> /eventos
Q3	M11 - Número de fluxos de controlo entre participantes
	ME7 - Tempo de fixação nos fluxos de controlo entre participantes

Continuação na próxima página

Tabela 4.1 – *Continuação da página anterior*

	ME8 - Percentagem do tempo de fixação em fluxos de controlo entre participantes
	ME9 - Número de movimentos sacadas entre participantes
Q4	M12 - Número de participantes M13 - Número de <i>Pools</i> ME10 - Percentagem do tempo de fixação em movimentos sacadas entre participantes ME9 - Número de movimentos sacadas entre participantes
Q5	M14 - Número de dados transferidos de cada atividade M15 - Número de dados gerados por atividade M16 - Número de dados transmitidos por atividade M17 - Número de dados M18 - Número de coleção de dados ME11 - Tempo de fixação em dados no modelo ME12 - Percentagem do tempo fixação em dados ME13 - Número de movimentos sacadas das associações dos dados ME14 - Percentagem de tempo em movimentos sacadas em associação de dados
Q6	M19 - Número de fluxos de controlo de saída && entrada M20 - Número de fluxos de sequencia M21 - Número de fluxos de mensagens M22 - Número de associações M23 - Média de fluxos de controlo de saída/entrada M24 - Máximo de fluxos de controlo de saída/entrada M25 - Mínimo de fluxos de controlo de saída/entrada ME15 - Tempo de fixação em fluxos de controlo de saída/entrada das atividades ME16 - Número de movimentos sacadas nos fluxos de controlo de saída/entrada das atividades
Q7	M33 - Número de fluxos paralelos M34 - Número de elementos que geram fluxos paralelos ME24 - Número de movimentos sacadas em fluxos paralelos ME25 - Percentagem do tempo em movimentos sacadas em fluxos paralelos
Q8	M32 - Número de ciclos ME22 - Tempo de fixação em ciclos ME23 - Percentagem do tempo de fixação em ciclos
Q9	M26 - Número de sub-processos M27 - Média dos elementos dos sub-processos

Continuação na próxima página

Tabela 4.1 – *Continuação da página anterior*

	M28 - Número mínimo de elementos dos sub-processos
	M29 - Número máximo de elementos dos sub-processos
	ME17 - Tempo de fixação dos sub-processos
	ME18 - Percentagem do tempo de fixação em sub-processos
	ME19 - Número de movimentos sacadas entre diferentes sub-processo
	ME20 - Percentagem de tempo em movimentos sacadas entre diferentes sub-processo
Q10	M30 - Número de elementos por sub-processo
	M31 - Número de elementos de cada género por sub-processo
	ME17 - Tempo de fixação dos sub-processos
	ME18 - Percentagem do tempo de fixação em sub-processos
	ME21 - Número de movimentos sacadas nos elementos à volta de um sub-processo
	ME20 - Percentagem do tempo em movimentos sacadas entre diferentes sub-processos
Q11	ME30 - Número de movimentos sacadas para a legenda
	ME31 - Percentagem do tempo de fixação na legenda

4.2 Meta-Modelo BPMN

Para que seja possível recolher as métricas corretamente, deve ser definido um meta-modelo, pois será este a definir as propriedades dos objetos e regras dos modelos. Esta definição irá permitir a consulta e pesquisas nos modelos. De modo a obter um meta-modelo completo recorreu-se às fontes [Cor14; OMG11]. No Anexo I pode ser consultado um excerto do meta-modelo utilizado para a recolha das métricas. Os elementos apresentados no meta-modelo como classes são apresentados com os seus símbolos no Anexo II e a suas especificações são descritas na Secção 2.2.

De modo a facilitar a compreensão do meta-modelo este foi dividido em diversas partes, sendo que de seguida serão apresentadas algumas das mais relevantes. No primeiro esquema, na Figura 4.2 é apresentado os elementos base de um modelo BPMN. Os elementos base de um modelo BPMN podem ser: *FlowElementsContainer*, *FlowElement*, *MessageFlow*, *ItemAwareElement* e *Artifacts*.

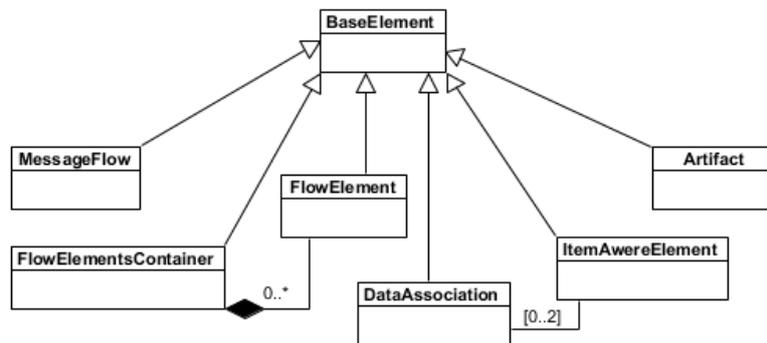


Figura 4.2: Raiz do meta-modelo BPMN adaptado de [Cor14]

Os *FlowElementsContainer* são os processos e os sub-processos de um modelo, tal como apresentado na Figura 4.3. Um processo pode conter vários processos e cada processo pode conter vários participantes. Mas tendo em conta que um participante só pode pertencer a um único processo.

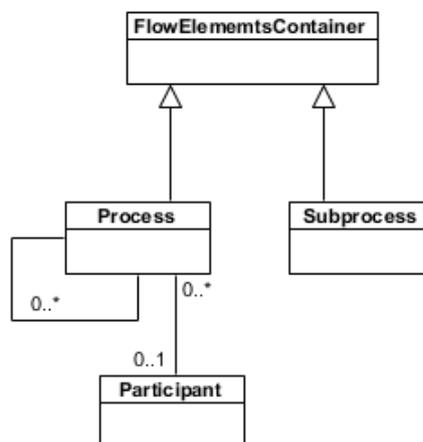


Figura 4.3: Excerto do meta-modelo BPMN correspondente à classe *FlowElementsContainer* adaptado de [Cor14]

Os *FlowElement*, Figura 4.4, representam os elementos das ações do modelo, ou por outras palavras, são compostos pelos *FlowNodes* que são a agregação das atividades, eventos e *gateways*. Por outro lado, os *SequenceFlow* são os fluxos de controlo do modelo e a ligação entre os diferentes *FlowNodes*. À parte destes dois elementos, existem os *DataObjects* estes são elementos de passagem de dados entre os diferentes *FlowNodes*.

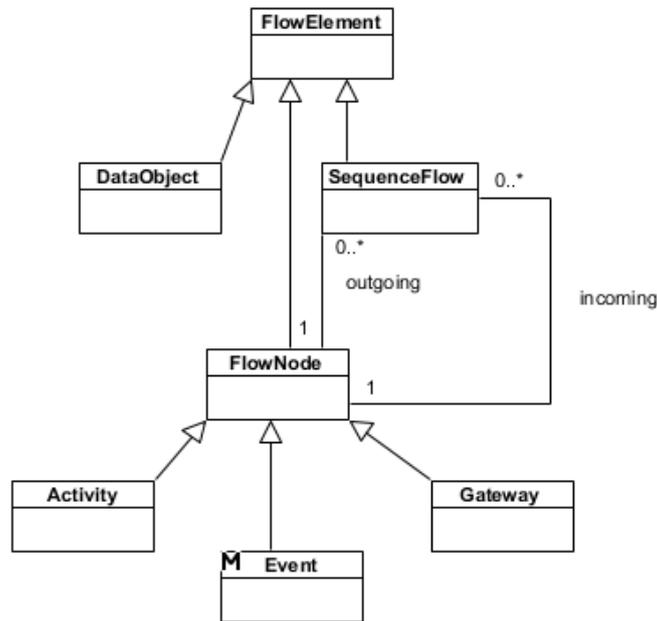


Figura 4.4: Excerto do meta-modelo BPMN correspondente à classe *FlowElement* adaptado de [Cor14]

Os *MessageFlow* são a ligação entre elementos relacionados entre si, estes elementos possuem a receção de uma mensagem ou o envio de uma mensagem, como é apresentado na Figura 4.5.

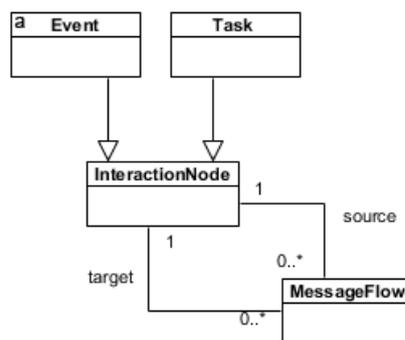


Figura 4.5: Excerto do meta-modelo BPMN correspondente à classe *MessageFlow* adaptado de [Cor14]

Os *ItemAwareElement*, Figura 4.6, consistem nos dados dos modelos e as suas respectivas associações. Cada dado gerado ou utilizado durante o processo modelado deve ter uma origem e um destino.

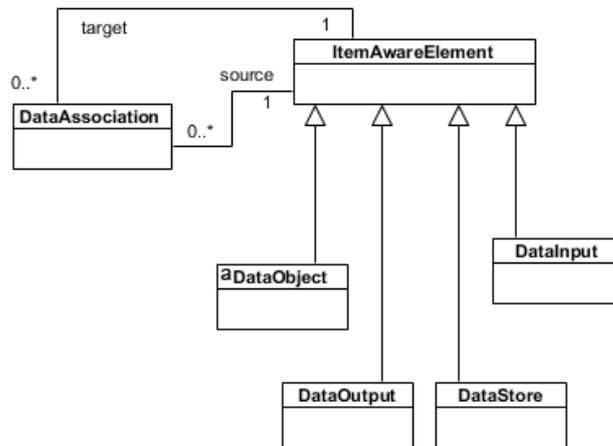


Figura 4.6: Excerto do meta-modelo BPMN correspondente à classe *ItemAwareElement* adaptado de [Cor14]

Os *Artifacts* servem para ajudar na compreensão dos modelos e para isso são utilizados os elementos: *Group*, *TextAnnotation* e *Association*, tal como demonstrado na Figura 4.7.

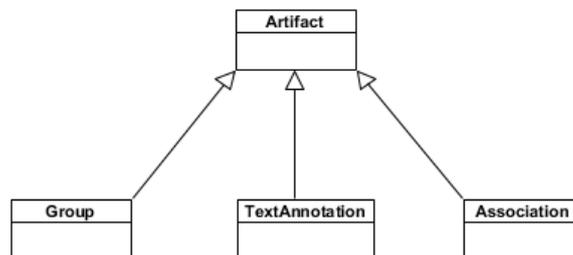


Figura 4.7: xcerto do meta-modelo BPMN correspondente à classe *Artifacts* adaptado de [Cor14]

4.3 Definição de métricas propostas

Com a especificação das métricas e a definição do meta-modelo tornou-se possível definir e formalizar as métricas apresentadas na tabela Tabela 4.1, quando estas dizem respeito aos modelos e ao seu respetivo *meta-modelo*. Para cada métricas novas é apresentada a sua definição rigorosa, em OCL, e uma definição informal, em linguagem natural. Para cada métrica é ainda apresentado qual o seu limite inferior e superior assim como um pequeno comentário caso haja necessidade para tal, para ajudar à compreensão da mesma.

Apesar de terem sido implementadas todas as métricas apresentadas, neste capítulo serão apresentadas apenas aquelas que não foram encontradas na nomenclatura e em

determinados casos será apresentado um exemplo de métricas semelhantes, como por exemplo, a percentagem dos diferentes eventos ou *gateways*.

Questão 1

O objetivo da métrica M2, apresentada na [Tabela 4.2](#), é representar a maneira de como obter a percentagem de *gateways* dentro de todos os *BaseElements*. Foi também especificada esta percentagem em relação aos *FlowElements* e *FlowNodes*.

Tabela 4.2: Métrica M2

Nome	M2 - Percentagem de <i>gateways</i>
Definição informal	Percentagem de <i>gateways</i> existente no modelo
Definição rigorosa	<pre> context Gateway::percentageGatewaysOnBaseElements() : Real pre prePercentageGatewayBaseElement: Gateway.allInstances()->select(a a.ocllsKindOf(Gateway)) -> size() > 0 and BaseElement.allInstances() -> size() > 0 post postPercentageGatewayBaseElement: result = Gateway.allInstances()->select(a a.ocllsKindOf(Gateway)) -> size() / BaseElement.allInstances() -> size() </pre>
Limite inferior	Ocorre quando existe apenas um elemento <i>gateway</i> no modelo
Limite Superior	Não existe um limite máximo de <i>gateways</i> para os modelos
Comentário	Existe a pré-condição de existir pelo menos um elemento <i>gateway</i> no modelo, pois o interesse desta, passa por perceber a capacidade de compreensão dos <i>gateways</i>

O objetivo da métrica M4, apresentada na [Tabela 4.3](#), é representar a maneira como é calculada a percentagem de um determinado género de *gateways*, neste caso, é apresentado como calcular a percentagem de *gateways* exclusivos dentro de todos os *gateways*.

Tabela 4.3: Métrica M4

Nome	M4 - Percentagem dos diferentes <i>gateways</i>
Definição informal	Percentagem de cada género de <i>gateways</i> existente no modelo
Definição rigorosa	<pre> context Gateway::percentageExclusiveGateways() : Real pre prePercentageExclusiveGateway: Gateway.allInstances()- >select(a a.ocllsKindOf(Gateway)) -> size() > 0 and Gateway.allInstances()->select(a a.ocllsKindOf(Exclusive- Gateway)) -> size() > 0 </pre>

Continuação na próxima página

Tabela 4.3 – Continuação da página anterior

	<pre> post postPercentageExclusiveGateway: result = Gateway.allInstances()->select(a a.ocIsKindOf(Exclu- siveGateway)) -> size() / Gateway.allInstances()->select(a a.ocIsKindOf(Gateway)) -> size() </pre>
Limite inferior	Ocorre quando existe apenas um elemento <i>gateway</i> no modelo
Limite Superior	Não existe um limite máximo de <i>gateways</i> para os modelos
Comentário	Existe a pré-condição de existir pelo menos um elemento <i>gateway</i> no modelo, pois o interesse desta, passa por perceber a capacidade de compreensão dos <i>gateways</i>

O objetivo da métrica M5, apresentada na Tabela 4.4 é a representação da soma de todos os géneros de *gateways* representados no modelo.

Tabela 4.4: Métrica M5

Nome	M5 - Número de <i>gateways</i> distintos
Definição informal	Diversidade de <i>gateways</i> existente no modelo
Definição rigorosa	<pre> context Gateway::numberDiferentGateways() : Integer pre preNumberDiferentGateways: Gateway.allInstances()- >select(a a.ocIsKindOf(Gateway)) -> size() > 0 post postNumberDiferentGateways: re- sult = Gateway.allInstances()->collect(a a.ocIsTypeOf(ExclusiveGateway))->asSet()- >count(true) + Gateway.allInstances()->collect(a a.ocIsTypeOf(InclusiveGateway))->asSet()- >count(true) + Gateway.allInstances()->collect(a a.ocIsTypeOf(ParallelGateway))->asSet()- >count(true) + Gateway.allInstances()->collect(a a.ocIsTypeOf(ComplexGateway))->asSet()- >count(true) + Gateway.allInstances()->collect(a a.ocIsTypeOf(EventBasedGateway))->asSet()->count(true) </pre>
Limite inferior	Ocorre quando existe apenas um elemento <i>gateway</i> no modelo
Limite Superior	Não existe um limite máximo de <i>gateways</i> para os modelos
Comentário	Existe a pré-condição de existir pelo menos um elemento <i>gateway</i> no modelo, pois o interesse desta, passa por perceber a capacidade de compreensão dos <i>gateways</i>

Questão 2

O objetivo da métrica M6, apresentada na [Tabela 4.5](#), é representar a maneira de como obter a percentagem de eventos dentro de todos os *BaseElements*. Foi também especificada esta percentagem em relação aos *FlowElements*, *FlowNodes* e *InteractionNodes*.

Tabela 4.5: Métrica M6

Nome	M6 - Percentagem de eventos
Definição informal	Percentagem de eventos existente no modelo
Definição rigorosa	<pre>context Event::percentageEventsOnBaseElements() : Real pre prePercentageEventBaseElement: Event.allInstances()- ->select(a a.ocIsKindOf(Event)) -> size() > 0 and BaseE- lement.allInstances() -> size() > 0 post postPercentageEventBaseElement: result = Event.allInstances()->select(a a.ocIsKindOf(Event)) -> size() / BaseElement.allInstances() -> size()</pre>
Limite inferior	Ocorre quando existe apenas um evento no modelo
Limite Superior	Não existe um limite máximo de eventos para os modelos
Comentário	Existe a pré-condição de existir pelo menos um evento no modelo, pois o interesse desta, passa por perceber a capacidade de compreensão dos eventos

O objetivo da métrica M7, apresentada na [Tabela 4.6](#), é o calculo da percentagem de um determinado género de evento, neste caso, é apresentado como calcular a percentagem de eventos de captura dentro de todos os eventos.

Tabela 4.6: Métrica M7

Nome	M7 - Percentagem dos diferentes eventos
Definição informal	Percentagem de cada género de eventos existente no modelo
Definição rigorosa	<pre>context Event::percentageCatchEvent() : Real pre prePercentageCatchEvent: Event.allInstances()->select(a a.ocIsKindOf(CatchEvent)) -> size() > 0 and Event.allInstances()->select(a a.ocIsKindOf(Event)) -> size() > 0 post postPercentageCatchEvent: result = Event.allInstances()- >select(a a.ocIsKindOf(CatchEvent)) -> size() / Event.allInstances()->select(a a.ocIsKindOf(Event)) -> size()</pre>
Limite inferior	Ocorre quando existe apenas um evento no modelo

Continuação na próxima página

Tabela 4.6 – Continuação da página anterior

Limite Superior	Não existe um limite máximo de eventos para os modelos
Comentário	Existe a pré-condição de existir pelo menos um evento no modelo, pois o interesse desta passa por perceber a capacidade de compreensão dos eventos

O objetivo da métrica M10, apresentada na [Tabela 4.7](#), é a representação da soma de todos os géneros de eventos representados no modelo.

Tabela 4.7: Métrica M10

Nome	M10 - Número de eventos distintos
Definição informal	Diversidade de eventos existente no modelo
Definição rigorosa	<pre>context Event::numberDiferentEvents() : Integer pre preNumberDiferentEvents: Event.allInstances()->select(a a.oclIsKindOf(Event)) -> size() > 0 post postNumberDiferentEvents: result = Event.allInstances()- >collect(a a.oclIsTypeOf(IntermediateCatchEvent))- >asSet()->count(true) + Event.allInstances()->collect(a a.oclIsTypeOf(StartEvent))->asSet()->count(true)+ Event.allInstances()->collect(a a.oclIsTypeOf(EndEvent))- >asSet()->count(true) + Event.allInstances()->collect(a a.oclIsTypeOf(ImplicitThrowEvent))->asSet()- >count(true) + Event.allInstances()->collect(a a.oclIsTypeOf(BoundaryEvent))->asSet()->count(true)</pre>
Limite inferior	Ocorre quando existe apenas um evento no modelo
Limite Superior	Não existe um limite máximo de eventos para os modelos
Comentário	Existe a pré-condição de existir pelo menos um evento no modelo, pois o interesse desta, passa por perceber a capacidade de compreensão dos eventos

Questão 5

O objetivo da métrica M15 apresentada na [Tabela 4.8](#), é representar a quantidade de elementos de dados que são gerados pelas atividades na sua execução.

Tabela 4.8: Métrica M15

Nome	M15 - Número de dados gerados por atividades
-------------	-----------------------------------------------------

Continuação na próxima página

Tabela 4.8 – Continuação da página anterior

Definição informal	Quantidade de dados produzidos pelas atividades dos processos no modelo
Definição rigorosa	<pre>context BaseElement::numberDataOutput() : Integer pre preNumberDataOutput: DataOutput.allInstances()->size()>0 post postNumberDataOutput: result = DataOutput.allInstances()->size() - Activity.allInstances().outgoing.targetRef</pre>
Limite inferior	Ocorre quando existe apenas um dado gerado no modelo
Limite Superior	Não existe um limite máximo de dados gerados para os modelos
Comentário	Existe a pré-condição de existir pelo menos uma atividade a gerar um determinado dado no modelo, pois o interesse desta, passa por perceber a capacidade de compreensão dos mesmos.

O objetivo da métrica M16, apresentada na [Tabela 4.9](#), é representar os dados inversos da tabela anterior, ou seja, representa os dados que são utilizados para a execução de uma atividade.

Tabela 4.9: Métrica M16

Nome	M16 - Número de dados transmitidos por atividades
Definição informal	Quantidade de dados recebidos pelas atividades dos processos no modelo
Definição rigorosa	<pre>context BaseElement::numberDataInput() : Integer pre preNumberDataInput: DataInput.allInstances()->size()>0 post postNumberDataInput: result = DataInput.allInstances()->size() - Activity.allInstances().incoming.targetRef</pre>
Limite inferior	Ocorre quando existe apenas um dado produzido no modelo
Limite Superior	Não existe um limite máximo de dados produzido para os modelos
Comentário	Existe a pré-condição de existir pelo menos uma atividade a receber um determinado dado no modelo, pois o interesse desta, passa por perceber a capacidade de compreensão dos mesmos.

O objetivo da métrica M17, apresentada na [Tabela 4.10](#), trata-se do somatório de todos os dados no modelo que sejam da classe objetos.

Tabela 4.10: Métrica M17

Nome	M17 - Número de dados
Definição informal	Quantidade de dados nos processos no modelo
Definição rigorosa	<pre>context BaseElement::numberDataObjects() : Integer pre preNumberDataObjects: DataObject.allInstances()- >size()>0 post postNumberDataObjects: result = DataObject.allInstances()->size()</pre>
Limite inferior	Ocorre quando existe apenas um dado no modelo
Limite Superior	Não existe um limite máximo de dados num modelos
Comentário	Existe a pré-condição de existir pelo menos uma atividade a receber ou a gerar um determinado dado no modelo, pois o interesse desta, passa por perceber a capacidade de compreensão dos mesmos.

O objetivo da métrica M18, apresentada na [Tabela 4.11](#), trata-se do somatório de todos os dados no modelo que sejam uma coleção de dados, ou seja, os elementos que tem uma função de base de dados no modelo.

Tabela 4.11: Métrica M18

Nome	M18 - Número de coleção de dados
Definição informal	Quantidade de coleção de dados nos processos no modelo
Definição rigorosa	<pre>context BaseElement::numberDataStore() : Integer pre preNumberDataStore: DataStore.allInstances()- >excluding(@DataStore)>0 post postNumberDataStore: result = DataStore.allInstances()- >excluding(@DataStore)->size()</pre>
Limite inferior	Ocorre quando existe apenas uma coleção de dado no modelo
Limite Superior	Não existe um limite máximo de coleção de dados num modelos
Comentário	Existe a pré-condição de existir pelo menos uma atividade a receber ou a gerar uma determinada coleção de dados no modelo, pois o interesse desta, passa por perceber a capacidade de compreensão dos mesmos.

Questão 6

O objetivo da métrica M19, apresentada na [Tabela 4.12](#), é representar todos os fluxos de entrada de cada uma das atividades, ou seja, o resultado é um conjunto de números e

cada um deles representa uma atividade específica e a respetiva quantidade de fluxos de entrada.

Tabela 4.12: Métrica M19

Nome	M19 - Número de fluxos de controlo de entrada
Definição informal	Quantidade de fluxos de controlo de entrada nos elementos
Definição rigorosa	<pre>context BaseElement::numberControlFlowEntry() : Bag(Integer) pre preNumberControlFlowEntry: Activity.allInstances()->size()>0 post postNumberControlFlowEntry: result = Activity.allInstances()->collect(a a.inputSequenceFlows()->size())</pre>
Limite inferior	Ocorre quando existe apenas um fluxo de controlo no modelo
Limite Superior	Não existe um limite máximo de fluxos de controlo no modelos
Comentário	Existe a pré-condição de existir pelo menos um fluxo de controlo no modelo, pois o interesse desta, passa por perceber a capacidade de compreensão dos mesmos.

O objetivo da métrica M23, apresentada na [Tabela 4.13](#), é o calculo da média dos fluxos de entrada de todas as atividades no modelo.

Tabela 4.13: Métrica M23

Nome	M23 - Média de fluxos de controlo de entrada
Definição informal	Média da quantidade de fluxos de controlo de entrada nos elementos
Definição rigorosa	<pre>context BaseElement::avgControlFlowEntry() : Integer pre preAvgControlFlowEntry: Activity.allInstances()->size()>0 post postAvgControlFlowEntry: result = Activity.allInstances()->collect(a a.inputSequenceFlows()->size())/Activity.allInstances()->size()</pre>
Limite inferior	Ocorre quando existe apenas um fluxo de controlo no modelo
Limite Superior	Não existe um limite máximo de fluxos de controlo no modelos
Comentário	Existe a pré-condição de existir pelo menos um fluxo de controlo no modelo, pois o interesse desta, passa por perceber a capacidade de compreensão dos mesmos.

O objetivo da métrica M24, apresentada na [Tabela 4.14](#), é representar o maior número

de fluxos de entrada que ocorreu nas atividades do modelo.

Tabela 4.14: Métrica M24

Nome	M24 - Máximo de fluxos de controlo de entrada
Definição informal	Máximo da quantidade de fluxos de controlo de entrada nos elementos
Definição rigorosa	<pre>context BaseElement::maxControlFlowEntry() : Integer pre preMaxControlFlowEntry: Activity.allInstances()->size()>0 post postMaxControlFlowEntry: result = Activity.allInstances()->iterate(e:Activity; max:Integer = -1 let ndg:Integer = e.inputSequenceFlows()->size() in if max = -1 then ndg else max.max(ndg) endif)</pre>
Limite inferior	Ocorre quando existe apenas um fluxo de controlo no modelo
Limite Superior	Não existe um limite máximo de fluxos de controlo no modelos
Comentário	Existe a pré-condição de existir pelo menos um fluxo de controlo no modelo, pois o interesse desta, passa por perceber a capacidade de compreensão dos mesmos.

O objetivo da métrica M25, apresentada na [Tabela 4.15](#), é representar o número mínimo de fluxos de entrada que ocorreu nas atividades do modelo.

Tabela 4.15: Métrica M25

Nome	M25 - Mínimo de fluxos de controlo de entrada
Definição informal	Mínimo da quantidade de fluxos de controlo de entrada nos elementos
Definição rigorosa	<pre>context BaseElement::minControlFlowEntry() : Integer pre preMinControlFlowEntry: Activity.allInstances()->size()>0 post postMinControlFlowEntry: result = Activity.allInstances()->iterate(e:Activity; min:Integer = -1 let ndg:Integer = e.inputSequenceFlows()->size() in if min = -1 then ndg else min.min(ndg) endif)</pre>
Limite inferior	Ocorre quando existe apenas um fluxo de controlo no modelo
Limite Superior	Não existe um limite máximo de fluxos de controlo no modelos
Comentário	Existe a pré-condição de existir pelo menos um fluxo de controlo no modelo, pois o interesse desta, passa por perceber a capacidade de compreensão dos mesmos.

Questão 7

O objetivo da métrica M33, apresentada na [Tabela 4.16](#), é obter o número de fluxos paralelos existentes num modelo, frequentemente criados devido à utilização de *gateways* com a função de divisão ou opção do fluxo de controlo.

Tabela 4.16: Métrica M33

Nome	M33 - Número de fluxos paralelos
Definição informal	Quantidade de fluxos paralelos num modelo
Definição rigorosa	<pre>context FlowNode::numberParallelFlux() : Integer pre preNumberParallelFlux: FlowNode.allInstances()->size()>0 post postNumberParallelFlux: result = (Gateway.allInstances()->select(a not a.oclIsKindOf(ExclusiveGateway) and not a.oclIsKindOf(ComplexGateway) and a.outgoing_a.targetRef->size())>1) ->union(Activity.allInstances()->select(a a.outgoing_a.targetRef->size())>1) ->union(Task.allInstances()->select(a a.outgoing_a.targetRef->size())>1) ->union(Event.allInstances()->select(a a.outgoing_a.targetRef->size())>1))) ->iterate(elem: FlowNode; acc : Bag(FlowNode) = Bag acc->union(elem.outgoing_a.targetRef)->size())</pre>
Limite inferior	Ocorre quando existe apenas dois fluxos paralelos no modelo
Limite Superior	Não existe um limite máximo de fluxos paralelos no modelos
Comentário	Existe a pré-condição de existir pelo menos dois fluxos paralelos no modelo, pois o interesse desta, passa por perceber a capacidade de compreensão dos mesmos.

O objetivo da métrica M34, apresentada na [Tabela 4.17](#), é obter a quantidade de fluxos de controlo paralelos que os elementos do modelo geram ou podem ativar.

Tabela 4.17: Métrica M34

Nome	M34 - Número de elementos que geram fluxos paralelos
Definição informal	Quantidade elementos que podem ativar múltiplos fluxos num modelo
Definição rigorosa	<pre>context FlowNode::ParallelFlux() : Bag(FlowNode) pre preParallelFlux: FlowNode.allInstances()->size()>0</pre>

Continuação na próxima página

Tabela 4.17 – Continuação da página anterior

	<pre> post postParellelFlux: result = (Gateway.allInstances()- ->select(a not a.oclIsKindOf(ExclusiveGateway) and not a.oclIsKindOf(ComplexGateway) and a.outgoing_a.targetRef- ->size()>1) ->union(Activity.allInstances()- ->select(a a.outgoing_a.targetRef->size()>1)) - ->union(Task.allInstances()->select(a a.outgoing_a.targetRef- ->size()>1)) ->union(Event.allInstances()->select(a a.outgoing_a.targetRef->size()>1))) ->iterate(elem: FlowNode; acc : Bag(FlowNode) = Bag acc- ->union(elem.outgoing_a.targetRef)) </pre>
Limite inferior	Ocorre quando existe apenas dois fluxos paralelos em um elemento do modelo
Limite Superior	Não existe um limite máximo de fluxos paralelos que um elemento pode ativar no modelo
Comentário	Existe a pré-condição de existir pelo menos dois fluxos paralelos no modelo, pois o interesse desta, passa por perceber a capacidade de compreensão dos mesmos.

Questão 8

O objetivo da métrica M32, apresentada na [Tabela 4.18](#), é obter o número de ciclos existentes num modelo.

Tabela 4.18: Métrica M32

Nome	M32 - Número de ciclos
Definição informal	Quantidade de ciclos existentes num modelo
Definição rigorosa	<pre> context FlowNode::numberLoop() : Integer pre preNumberLoop: FlowNode.allInstances()->size()>0 post postNumberLoop: result = FlowNode.allInstances()- ->collectNested(a FlowNode.allInstances()->iterate(elem : FlowNode; acc : Bag(FlowNode) = Bag acc- ->union(elem.outgoing_a.targetRef))->count(a)>1)- ->count(true) </pre>
Limite inferior	Ocorre quando existe apenas um ciclo no modelo
Limite Superior	Não existe um limite máximo de ciclos no modelos
Comentário	Existe a pré-condição de existir pelo menos um ciclo no modelo, pois o interesse desta, passa por perceber a capacidade de compreensão dos mesmos.

Questão 9

O objetivo da métrica M27, apresentada na [Tabela 4.19](#), é representar a média de elementos que os sub-processos contêm.

Tabela 4.19: Métrica M27

Nome	M27 - Média dos elementos dos sub-processos
Definição informal	Média da quantidade de elementos dos sub-processos de um modelo
Definição rigorosa	<pre>context FlowElementsContainer::avgSubProcessElements() : Integer pre preAvgSubProcessElements: SubProcess.allInstances()->size()>0 post postAvgSubProcessElements: result = SubProcess.allInstances()->collect(a a.flowElements)->size()/SubProcess.allInstances()->size()</pre>
Limite inferior	Ocorre quando existe apenas um sub-processo no modelo
Limite Superior	Não existe um limite máximo de sub-processos no modelos
Comentário	Existe a pré-condição de existir pelo menos um sub-processo no modelo, pois o interesse desta, passa por perceber a capacidade de compreensão dos mesmos.

O objetivo da métrica M28, apresentada na [Tabela 4.20](#), é representar o valor mínimo de elementos de todos os sub-processos do modelo.

Tabela 4.20: Métrica M28

Nome	M28 - Mínimo dos elementos dos sub-processos
Definição informal	Mínimo de elementos dos sub-processos de um modelo
Definição rigorosa	<pre>context FlowElementsContainer::minSubProcessElements() : Integer pre preMinSubProcessElements: SubProcess.allInstances()->size()>0 post postMinSubProcessElements: result = SubProcess.allInstances()->iterate(e:SubProcess; min:Integer = -1 let ndg:Integer = e.flowElements->size() in if min = -1 then ndg else min.min(ndg) endif)</pre>
Limite inferior	Ocorre quando existe apenas um sub-processo no modelo
Limite Superior	Não existe um limite máximo de sub-processos no modelos

Continuação na próxima página

Tabela 4.20 – *Continuação da página anterior*

Comentário	Existe a pré-condição de existir pelo menos um sub-processo no modelo, pois o interesse desta, passa por perceber a capacidade de compreensão dos mesmos.
-------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------

O objetivo da métrica M29, apresentada na [Tabela 4.21](#), trata-se do calculo do maior número de elementos existente nos diferentes sub-processos dos modelos.

Tabela 4.21: Métrica M29

Nome	M29 - Máximo dos elementos dos sub-processos
Definição informal	Máximo de elementos dos sub-processos de um modelo
Definição rigorosa	<pre>context FlowElementsContainer::maxSubProcessElements() : Integer pre preMaxSubProcessElements: SubProcess.allInstances()->size()>0 post postMaxSubProcessElements: result = SubProcess.allInstances()->iterate(e:SubProcess; max:Integer = -1 let ndg:Integer = e.flowElements->size() in if max = -1 then ndg else max.max(ndg) endif)</pre>
Limite inferior	Ocorre quando existe apenas um sub-processo no modelo
Limite Superior	Não existe um limite máximo de sub-processos no modelos
Comentário	Existe a pré-condição de existir pelo menos um sub-processo no modelo, pois o interesse desta, passa por perceber a capacidade de compreensão dos mesmos.

Questão 10

O objetivo da métrica M30, apresentada na [Tabela 4.22](#) é calcular o número de elementos de todos os sub-processos do modelo. O resultado desta métrica, será um conjunto de números que serão a representação do número de elementos de um sub-processo.

Tabela 4.22: Métrica M30

Nome	M30 - Número de elementos por sub-processos
Definição informal	Quantidade de elementos de cada um dos sub-processos de um modelo
Definição rigorosa	<pre>context FlowElementsContainer::numberElementsSubProcess() : Set(Integer) pre preNumberElementsSubProcess: SubProcess.allInstances()->size()>0</pre>

Continuação na próxima página

Tabela 4.22 – Continuação da página anterior

	<pre> post postNumberElementsSubProcess: result = SubProcess.allInstances()->collectNested(a a.flowElements- >size()) </pre>
Limite inferior	Ocorre quando existe apenas um sub-processo no modelo
Limite Superior	Não existe um limite máximo de sub-processos no modelos
Comentário	Existe a pré-condição de existir pelo menos um sub-processo no modelo, pois o interesse desta, passa por perceber a capacidade de compreensão dos mesmos.

O objetivo da métrica M31, apresentada na [Tabela 4.23](#), é apresentar a ocorrência de um determinado tipo de elemento nos sub-processos existentes no modelo. À semelhança das métricas dos *gateways* e eventos, esta métrica apresenta um exemplo de como calcular a quantidade de eventos nos sub-processos, mas o mesmo processo é repetido para todos os géneros de elementos dos modelos.

Tabela 4.23: Métrica M31

Nome	M31 - Número de elementos de cada género por sub-processos
Definição informal	Quantidade de um determinado elemento em cada um dos sub-processos de um modelo
Definição rigorosa	<pre> context FlowElementsContainer::numberEventsSubProcess() : Set(Integer) pre preNumberEventsSubProcess: SubProcess.allInstances()- >size()>0 post postNumberEventsSubProcess: result = SubProcess.allInstances()->collectNested(a a.flowElements- >select(b b.oclIsTypeOf(Event))->size()) </pre>
Limite inferior	Ocorre quando existe apenas um sub-processo no modelo
Limite Superior	Não existe um limite máximo de sub-processos no modelos
Comentário	Existe a pré-condição de existir pelo menos um sub-processo no modelo, pois o interesse desta, passa por perceber a capacidade de compreensão dos mesmos.

4.4 Sumário

De uma forma sucinta, este capítulo apresenta uma proposta para avaliar a complexidade dos modelos BPMN, de modo a responder e atingir o objetivo final desta dissertação.

IMPLEMENTAÇÃO DO PROTÓTIPO

5.1 Opções de Implementação para recolha de métricas

Existem quatro alternativas diferentes para implementar e desenvolver o *framework* para a recolha e análise dos modelos recorrendo às métricas em OCL. A primeira alternativa consiste na implementação de todo o *framework* de raiz, com recurso a linguagens de alto nível.

A segunda alternativa consiste na extensão de uma ferramenta já existente, na qual a funcionalidade a adicionar consistiria na criação de um *framework* de recolha de métricas. A análise das ferramentas existentes seria demasiado morosa, mas existe já alguns trabalhos que avaliam a validade e completude destas [Bar+15; Cor14; Fer17]. A ferramenta encontrada com as melhores características foi a **Bonita** [Bar+15; Fer17]. Para esta alternativa ser uma vantagem é necessário que permita recolher todas as métricas necessárias para a avaliação dos modelos.

A terceira alternativa recorre à construção de linguagens de domínio específico. Esta permite criar um *framework* não só para avaliar mas também para modelar, fornecendo *feedback* consoante as regras específicas especificadas. A grande vantagem é a automatização do processo de recolha de métricas. As tecnologias a serem utilizadas por esta são oferecidas pelo **IDE Eclipse Modeling Tools** [Alm14].

A quarta e última alternativa baseia-se no reaproveitamento do trabalho [Bar+15]. Ou seja, a utilização de uma ferramenta que permite não só importar modelos da ferramenta **bizagi** (a mais utilizada para modelar os modelos BPMN) e implementar novas métricas em OCL, apresentadas no Capítulo 4. A utilização desta ferramenta possui ainda a vantagem de já ter implementadas as métricas de modelação destes modelos, que são utilizadas para responder à pergunta Q11 e o meta-modelo utilizado recorre a documentação oficial,

fazendo dele uma base fiável. Para que seja possível recolher as métricas relativas ao *eye-tracking*, pode-se recorrer a qualquer ferramenta que permita apresentar o modelo a ser avaliado ou modificado. Para facilitar a interação entre os especialistas pode-se recorrer à plataforma **Bizagi Modeler**, devido a esta ser bastante utilizada.

Devido ao tempo disponível para a realização desta dissertação e o facto da quarta alternativa ter já implementadas as métricas de modelação, esta torna-se a alternativa mais viável e por isso é a metodologia seguida para proceder à recolha das métricas dos modelos.

5.2 Tecnologias utilizadas

As tecnologias utilizadas para a implementação do *framework* de avaliação dos modelos BPMN podem ser divididas em duas partes.

A primeira parte consiste na implementação de métricas que permite analisar os modelos de um modo estatístico. Para tal acontecer, foram implementadas novas métricas numa ferramenta previamente desenvolvida noutra dissertação de mestrado [Bar+15]. De uma maneira sucinta, esta ferramenta permite obter o resultado de todas as métricas que estejam implementadas no meta-modelo para o modelo pretendido.

Antes de mais é preciso ver esta ferramenta como um decodificador de ficheiros no formato *XML Process Definition Language (XPDL)* em objetos, implementada utilizando a linguagem de programação **Java**. Com a perceção dos elementos que existem no modelo é então possível aplicar as métricas existentes no meta-modelo. O meta-modelo implementado com recurso ao **OCL**, permite especificar quais os objetos existentes no modelo e quais as pré-condições e pós-condições das funções e associações dos modelos.

Para que esta ferramenta seja utilizada fora do contexto local onde foi desenvolvida a ferramenta, ou seja, foi optado pelo desenvolvimento de *webservices* REST cujos *endpoints* permitem obter os resultados das métricas de um dado modelo a partir de qualquer local.

O *framework* de avaliação dos modelos BPMN trata-se de um *web-site* onde é possível receber os dados estatísticos de um modelo, recorrendo a um *webservice* REST para aceder à ferramenta de processamento dos modelos em formato **XPDL**. A comunicação e obtenção dos resultados é feita via **AJAX**.

Virando o foco agora para o *web-site*, este segue a estrutura típica, ou seja, existe uma divisão entre a parte visual e a parte lógica. Para a parte visual ou por outras palavras o *front-end*, foi utilizado o **REACT**. Este trata-se de uma biblioteca especializada para construir interfaces, sendo a implementação feita com recurso à linguagem de programação *JavaScript* [Jav] e ao **HTML** [Htm]. De modo a tornar a interface mais apelativa e intuitiva, foi utilizado o **Materialize** [Mat] como *framework* CSS. A parte lógica, designa-se por *back-end* e foi implementada com o *framework* **Spring**. Esta *framework* consiste em um mecanismo de alto nível para modelar todo o conteúdo lógico necessário e desejado para um programa ou *web-site*. Neste caso em específico, o principal objetivo consistia em modelar e armazenar todos os dados dos utilizadores da plataforma assim como os

dados referentes aos testes para a validação empírica. A base de dados escolhida foi o *mysql*, sendo esta uma base de dados relacional com boa documentação e com uma boa implementação no que toca à sua conexão com o *Spring*.

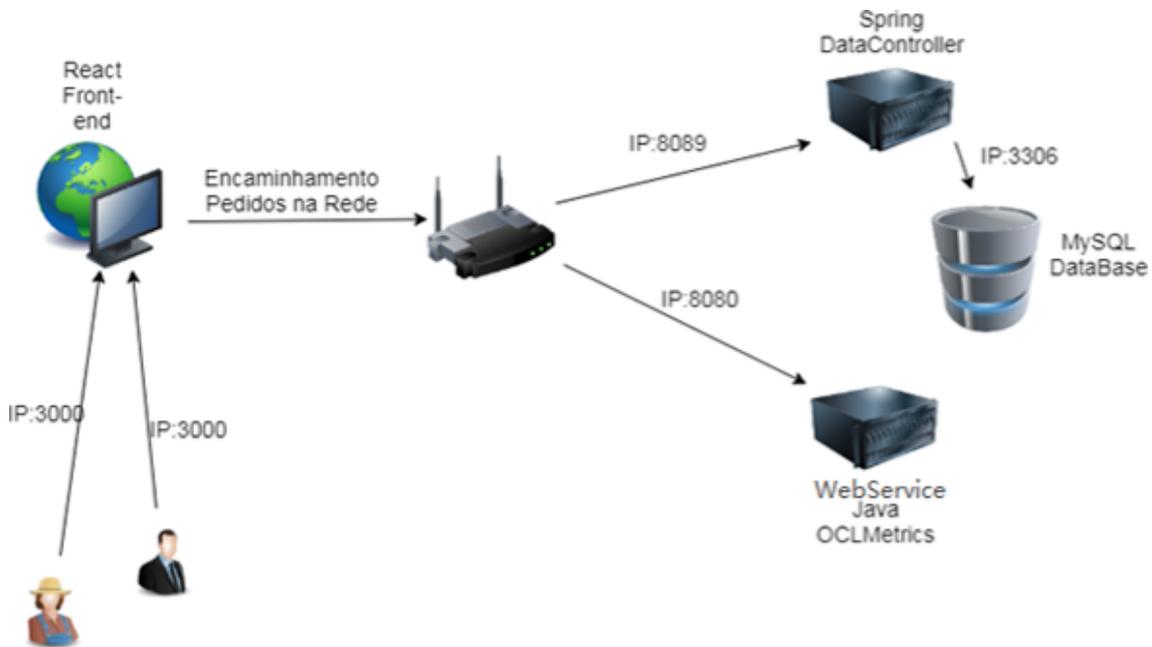


Figura 5.1: Esquema da arquitetura do protótipo

Na Figura 5.1 é apresentado um esquema da arquitetura geral do *framework* de avaliação desenvolvido. Este é composto por uma componente *front-end*, que permite a interação do utilizador com o código de aplicação das métricas. Esta é acedida através de um endereço IP, que por sua vez irá interagir através de um *webservice* com as componentes de *back-end* que irão fornecer a informação da aplicação das métricas no modelo *BPMN* submetido e com os testes da validação empírica. No que toca às comunicações entre as diferentes componentes, foi recorrido a pedidos assíncronos *Rest*.

Por ultimo, as componentes de *back-end* irão interagir com uma base de dados, que irá guardar e fornecer os dados necessários para a realização das funções acionadas.

5.3 Funcionalidades do protótipo

Devido ao facto da aplicação se encontrar disponível online em <http://85.243.14.201:3000>, esta pode ser acedida e utilizada por qualquer tipo de utilizador em qualquer tipo de computador e não existe nenhuma barreira causada pelo sistema operativo.

O *framework* de avaliação dos modelos *BPMN*, possui várias funcionalidades, podendo estas serem agrupadas em dois grupos.

O primeiro grupo de funcionalidades diz respeito às funcionalidades não específicas da plataforma, sendo elas:

- *Login* na plataforma.
- Registo na plataforma.

Por outro lado, o segundo grupo representa as funcionalidades específicas da plataforma, sendo elas:

- Submissão de um modelo no formato [XPDL](#) e obter os resultados da aplicação das métricas do respetivo modelo.
- Realização dos testes para a validação empírica.

Na plataforma é ainda possível visualizar uma apresentação sobre os modelos [BPMN](#), referente à sua simbologia e área de utilização.

No [Apêndice G](#) é apresentado um pequeno manual do utilizador, apresentado todos os procedimentos necessários para executar as funcionalidades apresentadas.

5.4 Sumário

Neste capítulo são apresentadas as alternativas de implementação encontradas para a implementada do protótipo desta dissertação e consequentemente qual foi a opção escolhida. É ainda apresentada uma descrição das tecnologias utilizadas para a implementação do protótipo e as funcionalidades que este apresenta.

VALIDAÇÃO TEÓRICA

Neste capítulo será apresentada a validação teórica das métricas implementadas que permitiram avaliar a complexidade dos modelos. A validação teórica apresentada tem como base as propriedades de *Weyuker* [Car05a].

6.1 Validação

As propriedades de *Weyuker* [Car05a], Tabela 6.1, são utilizadas para validar as métricas desenvolvidas com o objetivo de medir a complexidade dos modelos. Sendo que esta é baseada em nove propriedades distintas, apresentadas sucintamente através de uma forma rigorosa, enquanto que a descrição em linguagem natural pode ser encontrada no Capítulo 2.

Na representação das propriedades de forma rigorosas as letras representam modelos, enquanto que o módulo de um modelo representa o resultado da métrica. Quando o módulo é composto por mais de um modelo, o resultado trata-se da composição dos resultados desses modelos.

Tabela 6.1: Propriedades de *Weyuker*

#	Descrição das propriedades de <i>Weyuker</i>
1	$\exists P, \exists Q : P \neq Q \wedge P \neq Q $
2	$\exists P, \exists Q : P \neq Q \wedge P = Q $
3	$\exists P, \exists Q : P \neq Q \wedge P \simeq Q $
4	$\exists P, \exists Q : P \equiv Q \simeq P \neq Q $
5	$\forall P, \forall Q : P \leq P; Q \wedge Q \leq P; Q $
6	$\exists P, \exists Q; \exists R : P \neq Q \wedge P = Q \wedge P; R \neq Q; R $

Continuação na próxima página

Tabela 6.1 – Continuação da página anterior

	$\exists P, \exists Q; \exists R : P \neq Q \wedge P = Q \wedge R; P \neq R; Q $
7	$\exists P, \exists Q : Q = \text{Permutao}(P) \wedge P \neq Q $
8	$\exists P, \exists Q : Q = \text{AlterarNomenclatura}(P) \wedge P \neq Q $
9	$\exists P, \exists Q : P + Q < P; Q $

6.1.1 Métricas dos Modelos

Para perceber de que modo as métricas são avaliadas será apresentado de seguida um exemplo para cada uma das propriedades de *Weyuker*.

A primeira propriedade diz-nos que para processos diferentes o resultado da aplicação por norma deve ser diferente. Neste âmbito se transpusermos para a linguagem da métrica M1 que nos dá o número de *gateways* de um modelo, podemos dizer que a primeira propriedade é verificada se dois modelos distintos possuem diferentes quantidades de *gateways* nos modelos. Num exemplo prático, se contarmos os *gateways* dos modelos em [Apêndice C](#), podemos verificar que a quantidade de *gateways* é variável.

A segunda propriedade por outro lado, diz-nos que a aplicação das métricas deve ser capaz de refletir o aumento da dificuldade e complexidade dos modelos. Por outras palavras, com o aumento da complexidade dos modelos o número de *gateways* nos modelos deve crescer, o que pode ser verificado ao longo dos três modelos dos diferentes grupos de teste.

Já a terceira propriedade, diz-nos que para modelos semelhantes a métrica deve apresentar resultados semelhantes. Em tradução para a métrica M1, esta deve apresentar valores semelhantes de quantidade de *gateways* em modelos diferentes. Os modelos fáceis do grupo A e B são um bom exemplo, pois o primeiro apresenta quatro *gateways* e o segundo por sua vez, apresenta cinco *gateways*, sendo que estes dois apresentam semelhanças no seu grau de dificuldade e pode ser feita uma comparação entre eles.

A quarta propriedade diz-nos que apesar de dois modelos apresentarem a mesma funcionalidade não implica que possuam a mesma complexidade. Ou por outras palavras, dois modelos que modelem o mesmo processo não têm que apresentar o mesmo grau de complexidade e detalhe e tal não acontece uma vez que as modelações dependem da sua área de aplicação final.

A quinta propriedade diz-nos que a junção de dois modelos será igual ou maior que a complexidade da soma dos dois modelos em separado. No caso da métrica M1, a quantidade de *gateways* da junção de dois modelos pode ser a simples soma dos *gateways*. Caso algum dos modelos se torne um sub-processo de outro, ou caso os modelos sejam adaptados para representar os dois modelos o número de *gateways* irá crescer pois serão acrescentadas novas decisões e fluxos ao processo modelado.

A sexta propriedade, ao contrário das restantes, já engloba mais que dois processos, e no caso em que sejam realizadas operações de agregação de modos dois a dois, o resultado

da aplicação das métricas irá variar ser variável. Podemos verificar num exemplo prático, por exemplo o modelo fácil e médio do grupo A apresenta treze *gateways* enquanto que a junção do modelo fácil e difícil apresenta vinte *gateways*.

A sétima propriedade diz-nos que a troca de posição dos elementos dos modelos não resultará com que a métrica obtenha o mesmo resultado. No caso da métrica M1, esta não se torna verdade, pois a quantidade de *gateways* não sofre alterações mesmo que a ordem pela qual a soma seja alterada.

A oitava propriedade será verificada em todas as métricas apresentadas, pois em nenhuma delas o nome afetará o resultado da aplicação das métricas, uma vez que estas são de cariz estatístico.

Por último, a nona propriedade diz-nos que a aplicação das métricas na junção dos modelos deve retornar valores diferentes quando aplicados diferentes géneros de junções. Esta só será aplicável quando na junção de modelos surjam novos elementos não presentes nos modelos originais, mas este caso não será verificado na aplicação das métricas propostas, uma vez que elas não adicionam novos elementos aos modelos.

Na [Tabela 6.2](#), são apresentadas de um modo sucinto os resultados de todas as métricas. A primeira coluna representa o identificador da métrica e as restantes colunas representam o resultado das métricas em todas as nove propriedades de *Weyuker*.

Tabela 6.2: Avaliação das Métricas dos Modelos

M	1	2	3	4	5	6	7	8	9
M1	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M2	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M3	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M4	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M5	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M6	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M7	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M8	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M9	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M10	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M11	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M12	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M13	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M14	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M15	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M16	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M17	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M18	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não

Continuação na próxima página

Tabela 6.2 – *Continuação da página anterior*

M19	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M20	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M21	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M22	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M23	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M24	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M25	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M26	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M27	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M28	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M29	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M30	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não
M31	Sim	Sim	Sim	Sim	Sim	Sim	Não	Sim	Não

Como é possível de ver na [Tabela 6.2](#), existem duas propriedades que nunca são verificadas (sete e nove), sendo que de resto todas as restantes são verificadas em todas as métricas apresentadas. O facto de as métricas não verificarem todas as propriedades de *weyker* pode ser visto como algo recorrente, uma vez que o conjunto de métricas avaliado já tem um tamanho expressivo.

Em particular, a sétima propriedade é uma métrica que só é válida quando o modelo sofre alterações (quantidade de elementos e não localização) devido à alteração de lógica de fluxo do processo, o que não acontece com as métricas apresentadas, uma vez que os modelos se mantêm inalterados, independentemente da métrica aplicada. Por outro lado, a nona propriedade não é verificada com uma simples aplicação de métricas de avaliação, pois o objetivo destas é extrair informação sobre os modelos individualmente e posteriormente comparar as diferentes informações. Por tal acontecer, a não validação desta propriedades em particular não se tornam relevantes.

6.1.2 Métricas do Eye-tracker

À semelhança das métricas dos modelos, na [Tabela 6.3](#), são apresentadas de um modo sucinto os resultados de todas as métricas. A primeira coluna representa o identificador da métrica e as restantes colunas representam o resultado da métricas em todas as nove propriedades de *Weyker*.

À semelhança das métricas dos modelos, para perceber o modo de como as métricas são avaliadas será apresentado uma breve explicação para cada uma das propriedades de *Weyker*.

No âmbito do *eye-tracker* a métrica ME1 dá-nos o tempo de fixação de uma pessoa nos *gateways* de um modelo. Podemos dizer que a primeira propriedade é verificada se dois

participantes distintos passarem tempos diferentes tempos de fixação nos *gateways* dos modelos.

No caso da segunda propriedade, o aumento da complexidade dos modelos no que toca ao seu número de *gateways*, influencia o tempo de fixação dos participantes nestes elementos.

Já a terceira propriedade, a métrica ME1, deve apresentar valores semelhantes de tempo de fixação em *gateways* em modelos diferentes. Esta métrica pode ser verificada caso o perfil dos participantes seja semelhante, uma vez que participantes com diferentes níveis de conhecimento irão enfrentar diferentes problemas ou pontos de concentração diferentes nos modelos.

Na quarta propriedade dois modelos que modelem o mesmo processo não têm que apresentar o mesmo grau de tempo de fixação nos *gateways*, esta é verificada pois cada modelo é um caso e irá ser diferente ao nível de interpretação por parte do participante.

Na aplicação da quinta propriedade, a validação da métrica ME1 será realizada tendo em conta a junção dos tempos de fixação nos *gateways*, ou seja, se com dois modelos distintos o tempo de fixação será igual ou maior que o tempo de fixação da soma dos dois modelos em separado. Esta propriedade será sempre verificada uma vez que haverá sempre a propriedade de comutatividade entre os diferentes modelos e estes nunca serão avaliados em simultâneo.

Quanto à sexta propriedade, o tempo de fixação irá variar entre todos os modelos e isto deve-se aos seus diferentes interpretadores, o que conseqüentemente provoca sempre diferenças nas somas dos tempos entre modelos.

No âmbito da sétima propriedade nas métricas do *eye-tracker*, podemos dizer que a troca de posição dos modelos nos testes com voluntários irá afetar o resultado das métricas. No caso da métrica ME1 o tempo de fixação nos *gateways* dos modelos irá depender do à vontade com a nomenclatura dos modelos e ao longo dos testes o participante começa a ter uma aprendizagem dos modelos e variará os seus tempos, e por sua vez a aprendizagem será influenciada pela ordem de visualização dos modelos.

A oitava propriedade será verificada em todas as métricas apresentadas, pois em nenhuma delas o nome afetará o resultado da aplicação das métricas, uma vez que estas são de cariz estatístico.

Por último, e à semelhança das métricas dos modelos, as métricas do *eye-tracker* não irão adicionar novos elementos aos modelos nem a sua junção irá produzir elementos diferentes daqueles que são existentes neles, por isso está não será validada como nas métricas anteriores.

Tabela 6.3: Avaliação das Métricas do Eye-tracker

M	1	2	3	4	5	6	7	8	9
ME1	Sim	Não							

Continuação na próxima página

Tabela 6.3 – *Continuação da página anterior*

ME2	Sim	Não							
ME3	Sim	Não							
ME4	Sim	Não							
ME5	Sim	Não							
ME6	Sim	Não							
ME7	Sim	Não							
ME8	Sim	Não							
ME9	Sim	Não							
ME10	Sim	Não							
ME11	Sim	Não							
ME12	Sim	Não							
ME13	Sim	Não							
ME14	Sim	Não							
ME15	Sim	Não							
ME16	Sim	Não							
ME17	Sim	Não							
ME18	Sim	Não							
ME19	Sim	Não							
ME20	Sim	Não							
ME21	Sim	Não							
ME22	Sim	Não							
ME23	Sim	Não							
ME24	Sim	Não							
ME25	Sim	Não							
ME26	Sim	Não							
ME27	Sim	Não							
ME28	Sim	Não							
ME29	Sim	Não							
ME30	Sim	Não							
ME31	Sim	Não							

Como é possível de ver na [Tabela 6.3](#), as métricas do *eye-tracker* só não são validadas na nona propriedade de *Weyuker*. Isto deve-se ao facto de que independentemente das métricas que são validadas, não serão adicionados novos elementos ao modelo, ou seja, as métricas relacionam-se com a compreensão do utilizador e não com os elementos dos modelos. Por isso, esta propriedade não ser validada não prejudica a validade das análises posteriores e a respetiva correção da aplicação das métricas nos modelos.

6.2 Sumário

Neste capítulo foi realizada a validação teórica das métricas implementadas e descritas no [Capítulo 4](#). Esta validação teve em conta as propriedades de *Weyker*.

VALIDAÇÃO EMPÍRICA

Neste capítulo são apresentadas as técnicas utilizadas para proceder a um estudo empírico relativo à compreensão dos modelos BPMN por parte de um determinado público alvo. Numa primeira parte, serão apresentados os casos de estudo para os testes do estudo empírico, sendo que posteriormente, são apresentadas todas as fases do plano de recolha de dados durante os testes realizados pelos voluntários do estudo.

7.1 Plano de experiências

Neste estudo empírico pretende-se extrair informações sobre a compreensão dos modelos por parte de um conjunto de voluntários. Para que seja possível recolher este género de dados é necessário recorrer a um dispositivo *eye-tracker* para posteriormente analisar os dados com recurso a ferramentas estatísticas. Neste estudo também será aplicado um conjunto de métricas aos dados recolhidos à semelhança do primeiro estudo. Estas por sua vez também são apresentadas no [Capítulo 4](#) e teoricamente validadas no [Capítulo 6](#).

Para tal nos seguintes tópicos é apresentado um plano com todos os detalhes para que seja possível realizar todas as experiências para a este estudo empírico.

Objetivos

O primeiro objetivo é perceber a compreensão dos modelos por parte dos seus utilizadores, mas devido a este ser demasiado genérico subdividiu-se em dez diferentes objetivos. Estes são relacionados com as perguntas apresentadas na abordagem GQM no [Capítulo 4](#).

- GE1 → compreender a dificuldade de compreensão dos *gateways*.
- GE2 → compreender a dificuldade de compreensão dos eventos.

- **GE3** → compreender a dificuldade de compreensão da interação entre participantes.
- **GE4** → compreender a dificuldade de compreensão dos participantes.
- **GE5** → compreender a dificuldade de compreensão dos dados do processo.
- **GE6** → compreender a dificuldade de compreensão dos fluxos de controlo das atividades.
- **GE7** → compreender a dificuldade de compreensão dos fluxos paralelos.
- **GE8** → compreender a dificuldade de compreensão dos ciclos.
- **GE9** → compreender a dificuldade de compreensão dos sub-processos.
- **GE10** → compreender a dificuldade de memorização da simbologia dos modelos.

Tarefas

De modo a conseguir alcançar os objetivos traçados foram desenhadas três tarefas distintas, sendo que todas elas seguem a mesma lógica de testes, alterando unicamente os modelos a testar e consequentemente as perguntas colocadas.

Antes de explicar os detalhes e diferenças entre os modelos de cada teste, será apresentada a estrutura geral de todos os testes. O início de cada teste começa com uma pequena explicação do contexto dos testes e posteriormente é pedido ao participante que leia a carta de consentimento, [Apêndice E](#). Após a leitura e concordância com os termos do teste o participante começa então a realização do seu teste.

Cada teste é composto por três modelos diferentes ([Apêndice C](#)), sendo que para cada modelo, o participante deve responder a cinco perguntas, não podendo voltar atrás na sua resposta. Após responder à quinta pergunta de cada modelo o participante deve responder a um pequeno questionário NASA-TLX, [Apêndice D](#), de modo a perceber a opinião do participante sobre o modelo analisado e as suas respetivas perguntas. É importante referir, que para cada um dos três modelos, as cinco perguntas que o participante tem que responder são sempre as mesmas, para efeitos de comparação e análise de dificuldade e compreensão.

Todas as perguntas a que o participante deve responder, possuem uma pequena lista de opções da qual o participante deve escolher recorrendo ao *rato*. Esta opção foi tomada para evitar que o utilizador desvie os olhos do ecrã durante o teste, deteriorando os resultados do *eye-tracker*.

No final de cada teste, o participante deve responder a um questionário demográfico. Tanto os modelos como as suas perguntas e o questionário demográfico podem ser consultadas no [Apêndice D e F](#).

Material

Como mencionado anteriormente, o material experimental que é necessário para proceder aos testes começa com a carta de consentimento do participante, os três modelos para cada um dos três grupos de teste, o que equivale a nove modelos distintos, o questionário dos modelos, a legenda genérica para todos os nove modelos, um questionário NASA-TLX e por último um questionário demográfico [Apêndice F](#).

Existem três géneros de testes, de modo a conseguir avaliar e atingir os objetivos definidos. Para avaliar a compreensão do modelo com o aumento da sua dificuldade, os três modelos de cada teste vão possuir diferentes graus de dificuldade, ou seja, existe um modelo categorizado como fácil, médio e difícil. Como existem três modelos para cada teste e de modo a mitigar o efeito de aprendizagem por parte dos participantes dos testes deve-se realizar testes com os modelos por ordens diferentes. Para perceber a divisão e ordenação dos modelos nos testes é apresentada na tabela [Tabela 7.1](#) a lógica da organização dos testes.

Tabela 7.1: Ordens dos níveis de dificuldade dos modelos nos grupos de testes

	O1	O2	O3	O4	O5	O6
Modelo Fácil	1º	1º	2º	3º	2º	3º
Modelo Médio	2º	3º	1º	1º	3º	2º
Modelo Difícil	3º	2º	3º	2º	1º	1º

O primeiro teste ou grupo A, tem como objetivo avaliar os participantes de um processo assim como os dados que nele são transmitidos. O grupo B pretende avaliar os sub-processos dos processos modelados. Por fim, o grupo C pretende avaliar os fluxos de controlo e os ciclos existentes nos processos. Para além dos grupos avaliarem elementos ou características específicas, estes irão também avaliar a capacidade de compreensão e memorização dos significados dos seus diferentes elementos. Assim como a memorização dos elementos, os dados relativos aos *gateways* e os eventos também serão avaliados em todos os testes.

Durante todas as experiências teve-se cuidado para não introduzir erros nem variáveis indesejáveis na execução dos testes. Devido a isso, todos os testes foram realizados no mesmo ecrã, *full hd* com vinte e duas polegadas e todos os modelos e questionários foram ajustados ao ecrã de modo a facilitar a leitura por parte do participante. É também importante frisar que todos os elementos utilizam a mesma simbologia e possuem o mesmo tamanho assim como o tamanho da letra das legendas dos elementos.

Por último falta referir que durante os testes não foi dada nenhuma descrição relativamente aos modelos nem foram apresentados a notação e significados dos elementos dos modelos [BPMN](#).

Participantes

Existem dois grupos de participantes distintos, o primeiro grupo pretende avaliar a compreensão dos modelos por parte de estudantes ou pessoas com pouca familiaridade com os modelos e por outro lado o segundo grupo pretende avaliar a compreensão dos modelos por parte de profissionais ou pessoas com alguma familiaridade com os BPMN. Independentemente do género de participantes, deve de haver um número semelhante de participantes entre os grupos e realizar os testes com todas as diferentes ordens possíveis, para que as comparações sejam o mais realista possível.

Para o primeiro grupo, ou seja, pessoas com pouco conhecimento dos modelos, foram realizadas sessenta experiências ao longo de duas semanas, sendo que destas cinco foram excluídas devido à baixa qualidade dos dados recolhidos pelo *eye-tracker*, resultando assim em cinquenta e cinco experiências elegíveis para análise.

De um modo mais detalhado, o grupo A teve a participação de 19 participantes, enquanto que o grupo B e C contou com 18 participantes cada.

Tabela 7.2: Número de voluntários sem conhecimento expressivo

Grupo A	Grupo B	Grupo C
19	18	18

Para o segundo grupo foram realizados poucos testes, mas apesar disso foram realizados 2 testes para o grupo A e B e 3 para o grupo C.

Tabela 7.3: Número de voluntários com conhecimento expressivo

Grupo A	Grupo B	Grupo C
2	2	3

Independentemente do grupo de participantes em que o voluntário seja enquadrado, nenhum repetiu nenhum teste ou procedeu à realização de mais do que um teste. Para cada participante foi recolhido um conjunto de dados demográficos após a conclusão dos testes como anteriormente referido, sendo estes:

- Idade
- Género
- Curso
- Nacionalidade
- Visão (sem acessórios de visão, com óculos ou com lentes de contacto)

- Ano do Curso
- Ocupação atual
- Conhecimento sobre os modelos [BPMN](#)

Caso exista algum conhecimento dos modelos, pretende-se saber ainda a duração da interação com os modelos e há quanto tempo foram utilizados.

Hipóteses, Parâmetros e Variáveis

As hipóteses de seguida apresentadas, são baseadas nas perguntas colocadas e cujo objetivo é conseguir obter uma resposta a elas através da enumeração de várias hipóteses e à adição de detalhe às mesmas. As hipóteses são compostas por duas premissas, a **H1** é a premissa positiva enquanto que a **H0** é a negação da premissa, ou seja, após a análise dos dados a premissa é verificada ou não. Pode haver casos em que os dados existentes não sejam o suficientemente expressivos para chegar a uma conclusão.

As hipóteses relacionadas com a perceção dos modelos por parte dos utilizadores dos modelos são:

Hipótese $H_{interacao}$

- **H1** - A interação com os modelos torna-se mais intuitiva com o aumento do tempo de interação com eles.
- **H0** - A interação com os modelos mantém-se inalterável com o aumento do tempo de interação com eles.

Hipótese $H_{confianca}$

- **H1** - O utilizador torna-se mais confiante à medida que começa a compreender os modelos e a sua lógica de navegação.
- **H0** - O utilizador fica mais frustrado devido à não compreensão os modelos e à sua lógica de navegação.

Hipótese $H_{prestacao}$

- **H1** - O utilizador consegue ter uma noção da sua prestação nos modelos.

- **H0** - O utilizador não consegue ter uma noção da sua prestação nos modelos.

As hipóteses relacionadas com os elementos específicos dos modelos são:

Hipótese $H_{QGateways}$

- **H1** - A grande quantidade de *gateways* nos modelos torna o modelo mais confuso e menos intuitivo no que toca à pesquisa no modelo por parte do utilizador.
- **H0** - A grande quantidade de *gateways* nos modelos não influencia a pesquisa no modelo por parte do utilizador.

Hipótese $H_{DGateways}$

- **H1** - A grande diversidade de *gateways* nos modelos torna incompreensível o papel de cada um dos *gateways* existentes.
- **H0** - A grande diversidade de *gateways* nos modelos é indiferente na compreensão do papel de cada um dos *gateways* existentes.

Hipótese $H_{FGateways}$

- **H1** - A funcionalidade dos diversos *gateways* é compreendida por parte do utilizador.
- **H0** - A funcionalidade dos diversos *gateways* não é compreendida por parte do utilizador.

Hipótese $H_{QEventos}$

- **H1** - A grande quantidade de eventos nos modelos torna o modelo mais confuso e menos intuitivo no que toca à pesquisa no modelo por parte do utilizador.

- **H0** - A grande quantidade de eventos nos modelos não influencia a pesquisa no modelo por parte do utilizador.

Hipótese $H_{DEventos}$

- **H1** - A grande diversidade de eventos nos modelos torna incompreensível o papel de cada um dos eventos existentes.
- **H0** - A grande diversidade de eventos nos modelos é indiferente na compreensão do papel de cada um dos eventos existentes.

Hipótese H_{FluxoC}

- **H1** - O fluxo de controlo e lógica do processo modelado é compreendido por parte do utilizador.
- **H0** - O fluxo de controlo e lógica do processo modelado não é compreendida por parte do utilizador.

Hipótese $H_{QParticipantes}$

- **H1** - A grande quantidade de participantes nos modelos torna o modelo mais confuso e menos intuitivo no que toca à pesquisa no modelo por parte do utilizador.
- **H0** - A grande quantidade de participantes nos modelos não influencia a pesquisa no modelo por parte do utilizador.

Hipótese $H_{PParticipante}$

- **H1** - É compreensível por parte do utilizador o papel e função de um participante nos modelos.

- **H0** - O papel de um participante num modelo é incompreensível parte do utilizador.

Hipótese $H_{Sub-processos}$

- **H1** - A existência de sub-processos é intuitiva por parte do utilizador.
- **H0** - A existência de sub-processos não é intuitiva por parte do utilizador.

Hipótese H_{Ciclos}

- **H1** - O utilizador é capaz de encontrar e compreender as atividades existentes nos ciclos dos processo.
- **H0** - O utilizador não compreende o conceito de ciclo nos modelos.

Hipótese $H_{Simbologia}$

- **H1** - A simbologia dos modelos é fácil de memorizar e intuitiva por parte do utilizador.
- **H0** - A simbologia dos modelos é difícil de memorizar e pouco intuitiva para parte do utilizador.

As hipóteses relacionadas com a duração dos testes por parte dos participantes no teste são:

Hipótese $H_{TempoResposta}$

- **H1** - Com o aumento da interação dos modelos o utilizador começa a diminuir o tempo de resposta às perguntas dos modelos.
- **H0** - Independentemente do tempo de interação dos modelos o utilizador mantém constante o tempo de resposta às perguntas dos modelos.

Hipótese $H_{TempoAssertividade}$

- **H1** - Um menor tempo de resposta de uma pergunta relaciona-se com o facto da resposta ser correta.
- **H0** - Um menor tempo de resposta de uma pergunta relaciona-se com o facto da resposta ser incorreta.

As hipóteses relacionadas com o conhecimento dos utilizadores sobre os modelos são:

Hipótese $H_{Padroes}$

- **H1** - Com o aumento da interação dos modelos o utilizador começa a ganhar comportamentos mais assertivos e padronizados na pesquisa nos modelos.
- **H0** - Independentemente do tempo de interação dos modelos o utilizador não ganha comportamentos mais assertivos e padronizados na pesquisa nos modelos.

Hipótese $H_{AfeitoAprendizagem}$

- **H1** - Com o aumento da interação dos modelos o utilizador começa a acertar mais perguntas devido ao efeito aprendizagem dos modelos.
- **H0** - Independentemente do tempo de interação dos modelos o utilizador mantém o mesmo comportamento de respostas às perguntas sobre os modelos.

Ambiente de Testes

Durante todos os testes teve-se o cuidado de fornecer aos participantes condições muito semelhantes para a realização dos seus testes. É preciso ter em conta que cada sessão de testes tinha unicamente um participante e tal como referido anteriormente antes de iniciar o teste foram dada ao participantes as instruções e indicações necessárias para o teste.

Em todos os testes o *eye-tracker* foi colocado debaixo do ecrã de teste e antes do participante ler a carta de consentimento e começar o seu teste, foram feitos ajustes ao

eye-tracker de modo a enquadrar o ângulo de visão do participante. Para além do ajuste do ângulo do *eye-tracker* é feita ainda uma calibração do dispositivo.

Procedimento de análise

Para todos os dados recolhidos foram coletados os seus dados estatísticos, tais como, o desvio padrões, derivações, *skewness* e *kurtosis*, isto de modo a obter uma visão mais detalhada das distribuições dos dados recolhidos. Após ter as estatísticas descritivas dos dados, serão aplicados testes de normalidade e consoante o resultado da normalidade dos dados serão feitas análises do efeito significativo dos dados. Este efeito é calculado através da aplicação dos testes do qui quadrado e do *oneway* entre grupos anova.

São também gerados gráficos, tais como *boxPlots* e *BarCharts*, para ter uma melhor perceção das distribuições dos dados.

Execução de testes

Nesta secção são detalhados todos os procedimentos e tecnologias usadas para proceder à realização dos testes com os voluntários.

Preparação

Os testes e a recolha dos dados foram realizados todos na mesma sala e com recurso a um ecrã de vinte e duas polegadas ligado a um computador portátil. Para além do ecrã é ainda necessário um dispositivo *eye-tracker* para a recolha de dados e por fim, um rato conectado ao computador.

Após o material montado, assim que o participante estivesse sentado de modo confortável na cadeira de testes, faz-se o ajuste da posição do *eye-tracker* em relação ao olhar do participante e no final, faz-se a calibração do dispositivo. Esta calibração corresponde ao participante seguir com o olhar um círculo no ecrã, permitindo assim o dispositivo perceber as características do olhar do participante em causa. O participante pode começar o seu teste assim que obtiver uma calibração qualificada como **boa** ou **excelente**, de modo a obter dados com um menor nível de desvios e erros de medição.

Desvios do plano

Alguns dos dados recolhidos não se encontravam de acordo com o esperado a nível de qualidade, uma vez que a recolha dos dados do olhar do participante não foi a adequada. Devido a isso em vez de ter sessenta experiências executadas com sucesso, foram aceites para análise cinquenta e cinco experiências, no que toca às experiências com voluntários com pouco conhecimento sobre os modelos.

No que toca aos testes com voluntários com algum conhecimento sobre os modelos não foram recolhidos dados suficientes até à data, mas aqueles obtidos cumpriram os requisitos de qualidade aceites para análise.

7.2 Sumário

Neste capítulo, foram abordados todos os detalhes relativos à preparação e execução dos estudos empíricos necessários para a análise e compreensão da complexidade dos modelos [BPMN](#). São ainda apresentadas às hipóteses as quais se pretende obter uma resposta com a análise dos resultados dos estudos.

ANÁLISE DOS DADOS RECOLHIDOS

Neste capítulo são apresentados todos os dados e resultados obtidos através das validações empíricas especificadas anteriormente. Depois de apresentados todos os dados e as suas respetivas hipóteses testadas, é apresentada uma discussão sobre os resultados obtidos assim como as ilações que podem ser retiradas dos dados.

8.1 Participantes

Os dados relativos aos participantes são os últimos dados da validação empírica a serem recolhidos. Estes são recolhidos após a conclusão dos testes para minimizar qualquer tipo de distração por parte do participante durante a realização dos testes e minimizar todos os impactos negativos na recolha dos dados com o *eye-tracker*.

Os dados recolhidos dos participantes foram apresentados em capítulos anteriores e podem ser consultados no [Apêndice F](#), sendo feito um resumo das suas informações neste capítulo. Na [Tabela 8.2](#) e [8.3](#) é apresentada uma análise descritiva dos participantes sem conhecimento ou com muito pouco conhecimento dos modelos.

O perfil dos participantes é composto pela frequência e percentagem das respostas dadas, tendo como base todas as hipóteses possíveis de resposta. Os dados dos participantes são divididos em duas tabelas, devido a certos dados serem obtidos através de uma resposta aberta ou fechada. Os dados de resposta fechada apresentam todas as hipóteses possíveis no questionário demográfico, enquanto que as perguntas de resposta aberta apresentam as opções dadas pelos participantes, tendo em atenção que respostas iguais mas com nomenclaturas diferentes foram agregadas.

Antes de apresentar as tabelas, é preciso ter em conta quantos participantes foram coletados para cada um dos dois perfis apresentados anteriormente. Estes valores são apresentados na [Tabela 8.1](#).

Tabela 8.1: Número de participantes

Sem conhecimento expressivo	Com conhecimento expressivo
55	7

Sem conhecimento expressivo

Tabela 8.2: Estatísticas descritivas dos participantes sem conhecimento expressivo (informação de resposta aberta)

	Dados Demográficos	Frequência	Percentagem
Idade	18	1	1,8%
	20	4	7,3%
	21	5	9,1%
	22	9	16,4%
	23	15	27,3%
	24	8	14,5%
	25	5	9,1%
	26	3	5,5%
	27	1	1,8%
	28	1	1,8%
	32	1	1,8%
	34	1	1,8%
	39	1	1,8%
Nacionalidade	Belga	1	1,8%
	Croata	1	1,8%
	Portuguesa	53	96,4%
Curso	Informática	55	100%

Tabela 8.3: Estatísticas descritivas dos participantes sem conhecimento expressivo (informação de resposta fechada)

	Possibilidades	Frequência	Percentagem
Género	F	12	21,8%
	M	43	78,2%
Visão	Sem acessórios	36	65,5%
	Com óculos	15	27,3%
	Com lentes	4	7,3%

Continuação na próxima página

Tabela 8.3 – *Continuação da página anterior*

Ano de Curso	1º Licenciatura	2	3,6%
	2º Licenciatura	1	1,8%
	3º Licenciatura	10	18,2%
	1º Mestrado	4	7,3%
	2º Mestrado	31	56,4%
	Doutoramento	4	7,3%
	Curso Finalizado	3	5,5%
Ocupação atual	Estudante	45	81,8%
	Investigador	3	5,5%
	Trabalhador	0	0%
	Trabalhador/Estudante	7	12,7%
Conhecimento	Sem Conhecimento	37	67,3%
	Conhecimento Académico	18	32,7%
	Conhecimento Profissional	0	0%
	Iniciativa Própria	0	0%

Com os dados apresentados nas tabelas acima, podemos afirmar que apesar dos participantes serem da área de informática e estarem maioritariamente no seu último ano de mestrado, não se reflete na existência de conhecimento destes modelos. Podemos afirmar que caso exista conhecimento sobre os modelos, este foi adquirido num ambiente académico em cadeiras de especialização da área da Engenharia de *Software*. Por outro lado, não existem dados que demonstrem como participantes de outras ocupações adquiriram os conhecimentos, devido à amostra dos participantes não ser muito diversificada na componente das suas ocupações.

Com conhecimento expressivo

Devido à quantidade pequena de testes realizados até ao momento neste grupo, estes dados não serão comparados nem utilizados nas análises apresentadas ao longo deste capítulo. Esta quantidade deve-se à dificuldade de contacto com profissionais na área e com ajustamento de marcação para os testes.

Esta lacuna é combatida com a quantidade expressiva de dados de participantes com pouco ou nenhum conhecimento sobre os modelos e o facto destes possuírem diferentes *backgrounds* na área de informática e modelação de processos.

8.2 Análises Estatísticas

Após a realização de todos os testes, pode-se proceder a uma análise estatística de modo a perceber a veracidade dos dados e o impacto real da dificuldade de interpretação dos modelos por parte dos utilizadores.

As análises são divididas pelos diferentes grupos, sendo estes o grupo A que teve por objetivo avaliar os participantes de um processo e os dados por eles transmitidos, o grupo B que teve por objetivo avaliar os sub-processos e por ultimo o grupo C que teve por objetivo avaliar os fluxos de controlo e os ciclos

8.2.1 Duração

Nesta secção serão apresentados os dados dos testes relativamente ao tempo de resposta às diferentes perguntas dos diferentes grupos e as suas respetivas análises estatísticas quando aplicáveis.

Dados

Os dados relativos aos tempos e respostas dadas consistem nos dados recolhidos durante os testes dos participantes relativos aos testes do estudo empírico com o *eye-tracker*. Estes ao contrário dos dados específicos dos participantes, são recolhidos durante toda a duração do teste. Para tal recolha ser possível, existe uma implementação da lógica da recolha dos dados no *framework* implementado. De uma maneira sucinta, cada vez que o participante passa à próxima pergunta ou modelo, é registada a resposta dada e o *timestamp* do momento da passagem da pergunta.

Em média os participantes sem conhecimento demoraram onze minutos a concluir os seus testes, como é possível ver na [Tabela 8.4](#) e na [Figura 8.1](#), sendo que o maior tempo registado em todos os testes de todos os grupos foi de vinte e cinco minutos, enquanto que o menor tempo registado foi de quatro minutos.

Tabela 8.4: Duração de todos os testes

	Duração Média	Duração Mínima	Duração Máxima
Sem conhecimento	11:23min	04:15min	25:17min

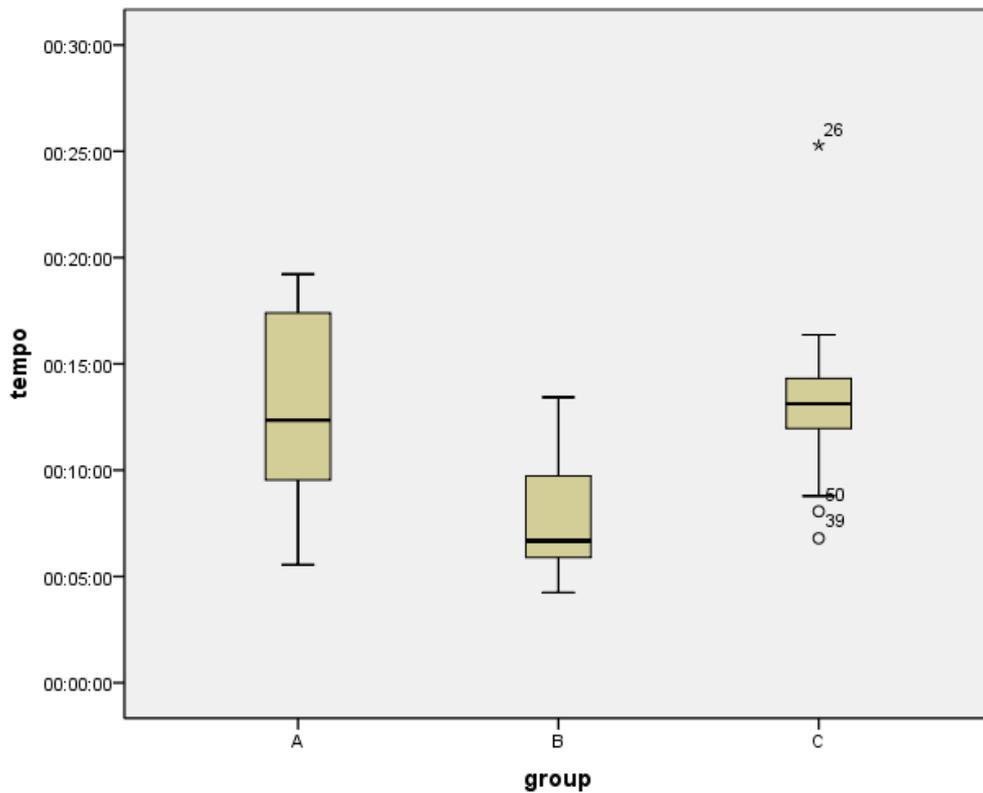


Figura 8.1: *BoxPlot* dos tempos dos participantes por grupos

Grupo A Na Tabela 8.5 e na Figura 8.2 é possível ver dados relativos aos tempos de realização dos testes do grupo A por parte dos participantes, sendo que os testes são separados pela posição em que é apresentado cada um dos três modelos dos testes.

Tabela 8.5: Estatísticas dos tempos do grupo A

Ordem	Duração Média	Modelo	Duração Média	Duração Mínima	Duração Máxima
O1	13:46min	Fácil	03:51min	03:00min	04:26min
		Médio	04:12min	03:50min	04:13min
		Difícil	04:24min	02:45min	05:54min
O2	17:00min	Fácil	05:25min	04:23min	06:15min
		Médio	03:05min	02:06min	04:01min
		Difícil	06:43min	05:52min	07:38min
O3	10:18min	Fácil	01:36min	01:31min	01:46min
		Médio	03:57min	03:17min	04:46min
		Difícil	02:49min	01:42min	04:11min
		Fácil	01:08min	00:46min	01:29min

Continuação na próxima página

Tabela 8.5 – Continuação da página anterior

O4	08:14min	Médio	04:00min	01:42min	04:31min
		Difícil	03:33min	01:10min	06:12min
		Fácil	03:18min	02:45min	03:46min
O5	17:38min	Médio	03:58min	03:09min	04:26min
		Difícil	07:57min	07:30min	08:41min
		Fácil	02:06min	01:00min	02:51min
O6	13:21min	Médio	03:47min	01:36min	05:27min
		Difícil	04:23min	03:22min	05:09min

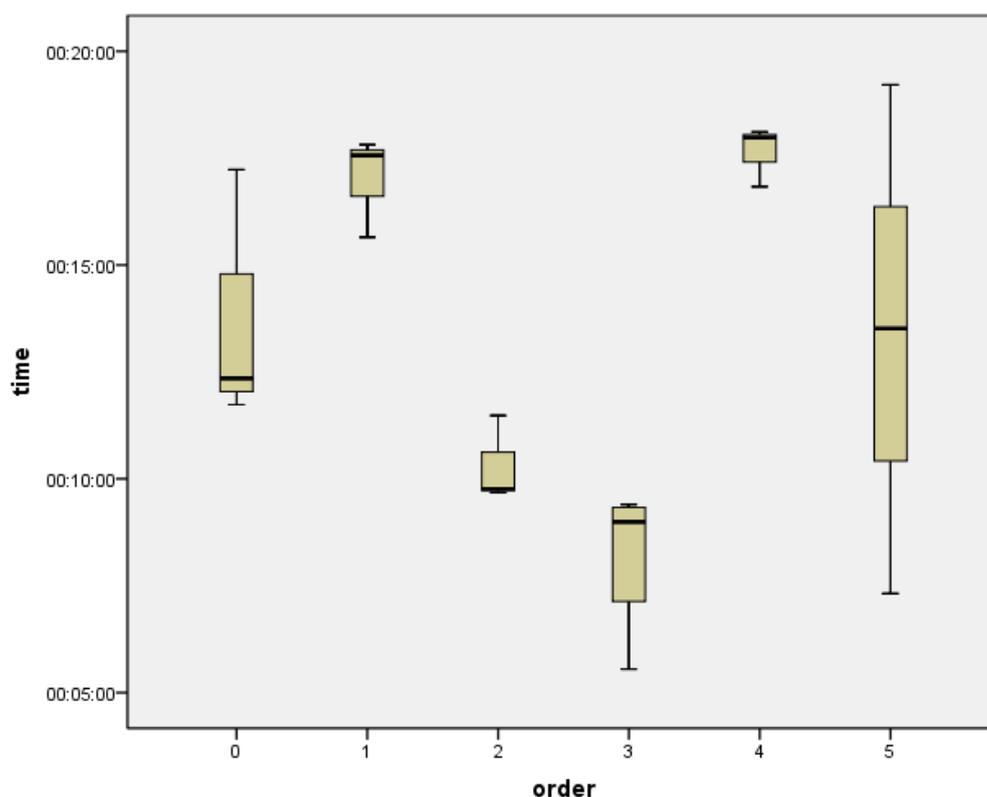


Figura 8.2: *BoxPlot* dos tempos dos participantes por ordem do grupo A

A duração média do grupo A, independentemente da sua ordem é de treze minutos e seis segundos. Ainda é possível visualizar nos dados que a duração dos testes varia consoante a ordem pela qual são apresentados os modelos e os seus diferentes níveis de dificuldade, algo que já era esperado uma vez que quando um modelo difícil é o primeiro a ser analisado o utilizador demora mais tempo a perceber e analisar o modelo.

Para além da duração do grupo A, temos ainda o tempo de resposta às cinco perguntas de cada modelo. Nas seguintes tabelas 8.6 e 8.4 e gráficos 8.3, 8.4 e 8.5 podemos observar

a duração média de cada pergunta relativamente à posição dos modelos no testes e à dificuldade do modelo.

Tabela 8.6: Estatísticas dos tempos das perguntas do grupo A

	Duração Média	Duração Mínima	Duração Máxima
Q1	01:03min	00:05min	03:28min
Q2	00:41min	00:02min	03:44min
Q3	01:05min	00:04min	04:14min
Q4	00:49min	00:03min	03:45min
Q5	00:51min	00:07min	04:18min

Tabela 8.7: Estatísticas dos tempos das perguntas dos modelos do grupo A

		Duração Média	Duração Mínima	Duração Máxima
modelo fácil	Q1	00:46min	00:05min	03:25min
	Q2	00:20min	00:04min	01:18min
	Q3	00:56min	00:04min	03:20min
	Q4	00:47min	00:03min	01:39min
	Q5	00:32min	00:08min	02:28min
modelo médio	Q1	00:59min	00:10min	01:38min
	Q2	00:41min	00:02min	03:44min
	Q3	01:08min	00:17min	04:18min
	Q4	00:45min	00:10min	01:34min
	Q5	00:42min	00:13min	01:47min
modelo difícil	Q1	01:24min	00:14min	03:28min
	Q2	01:02min	00:06min	03:38min
	Q3	01:12min	00:09min	03:00min
	Q4	00:56min	00:07min	03:45min
	Q5	01:17min	00:07min	04:18min

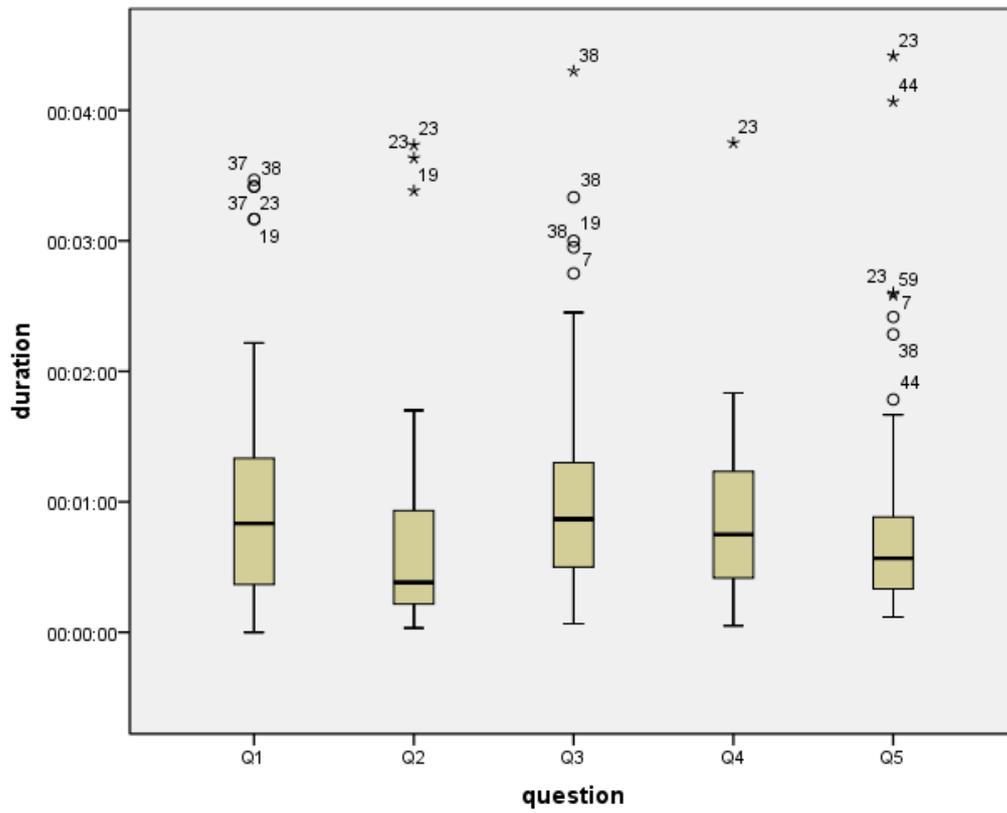


Figura 8.3: *BoxPlot* dos tempos dos participantes por pergunta do grupo A

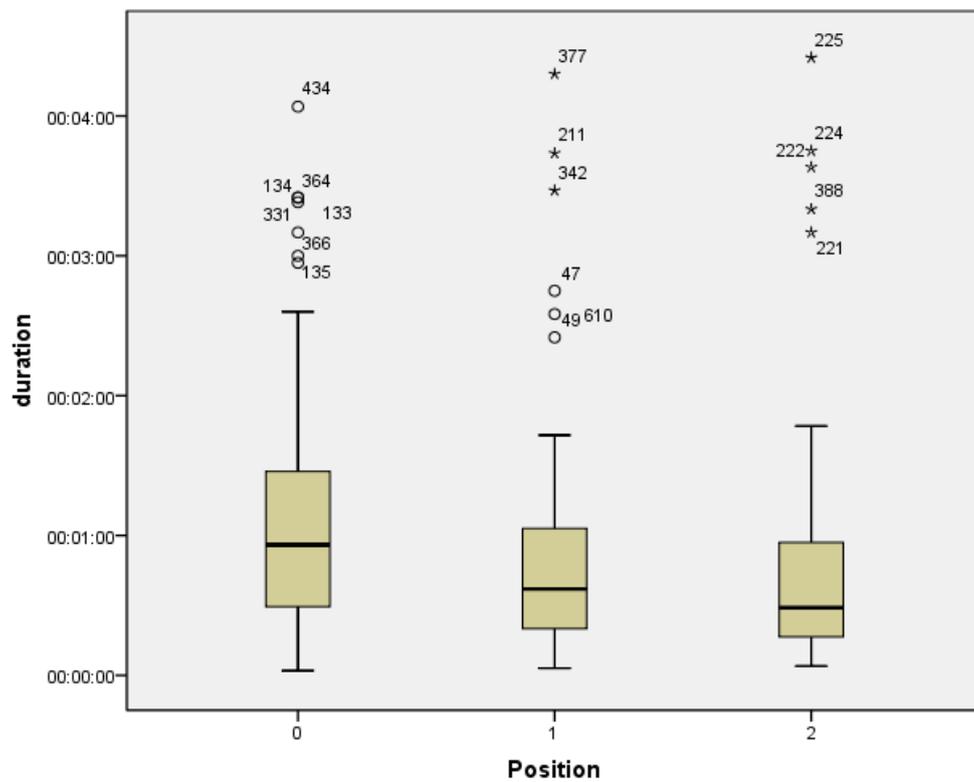


Figura 8.4: *BoxPlot* dos tempos dos participantes por posição do grupo A

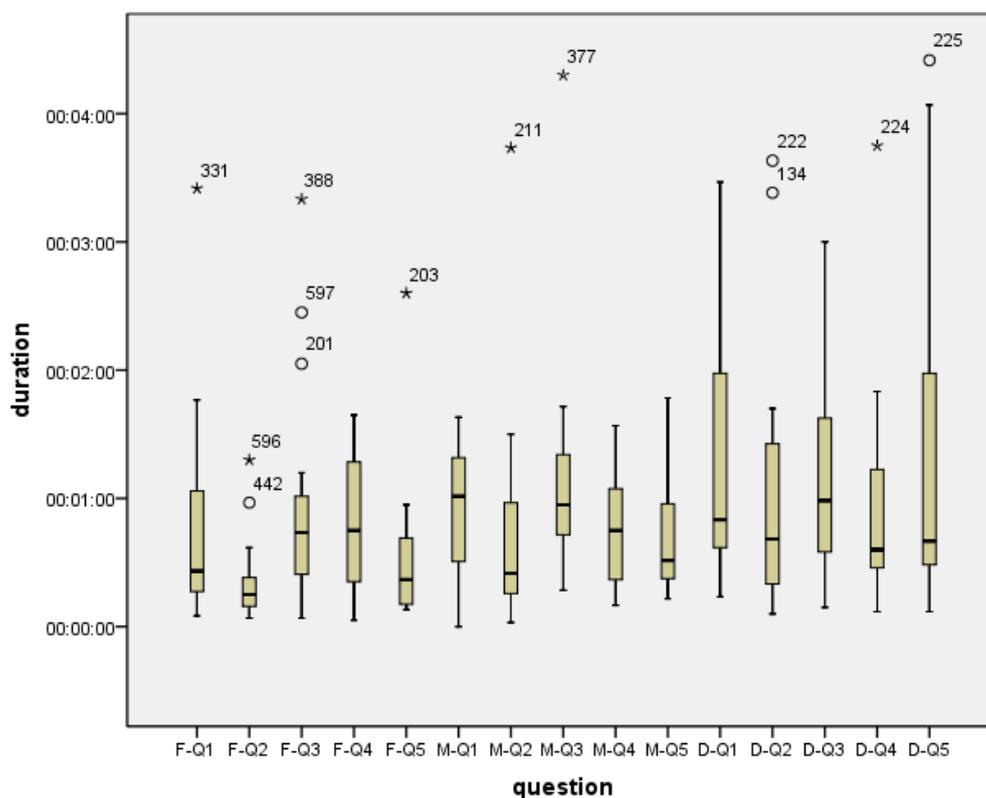


Figura 8.5: *BoxPlot* dos tempos dos participantes por dificuldade do grupo A

Grupo B Na Tabela 8.8 e na Figura 8.6 é possível ver dados relativos aos tempos de realização dos testes do grupo B por parte dos participantes, sendo que os testes são separados pela posição em que é apresentado cada um dos três modelos dos testes.

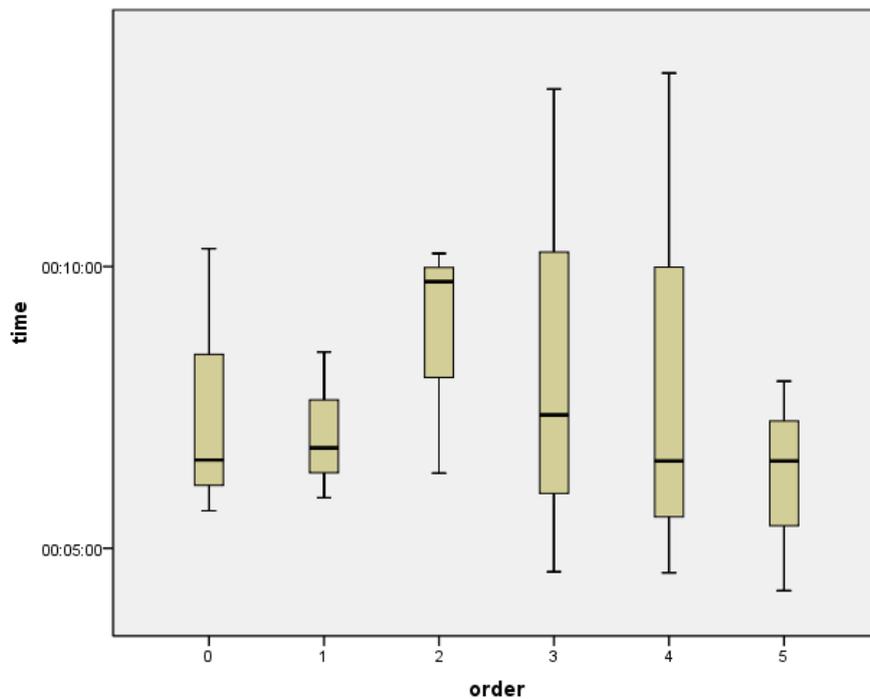
Tabela 8.8: Estatísticas descritivas dos tempos do grupo B

Ordem	Duração Média	Modelo	Duração Média	Duração Mínima	Duração Máxima
O1	07:31min	Fácil	01:26min	01:00min	02:01min
		Médio	02:20min	02:07min	02:47min
		Difícil	02:07min	01:18min	02:59min
O2	07:03min	Fácil	01:51min	01:23min	02:38min
		Médio	01:42min	01:35min	01:53min
		Difícil	02:32min	03:13min	01:28min
O3	08:46min	Fácil	01:01min	00:29min	02:04min
		Médio	03:30min	01:33min	04:27min
		Difícil	02:44min	01:33min	03:23min
		Fácil	00:49min	00:37min	01:16min

Continuação na próxima página

Tabela 8.8 – Continuação da página anterior

O4	08:22min	Médio	03:19min	01:39min	05:46min
		Difícil	02:32min	01:22min	03:36min
		Fácil	01:28min	00:41min	02:19min
O5	08:12min	Médio	01:48min	01:20min	02:22min
		Difícil	03:41min	01:36min	07:12min
		Fácil	00:31min	00:30min	00:33min
O6	06:15min	Médio	01:38min	01:23min	01:59min
		Difícil	02:48min	01:21min	04:12min

Figura 8.6: *BoxPlot* dos tempos dos participantes por ordem do grupo B

De um modo geral, o grupo B demora sete minutos e quarenta e um segundos a ser finalizado. Tal como era esperado, os modelos com mais duração de interpretação são os modelos médio e difícil independentemente da sua posição no teste.

Para além da duração do grupo B, temos ainda o tempo de resposta às cinco perguntas de cada modelo. Nas seguintes tabelas 8.9 e 8.7 e gráficos 8.7, 8.8 e 8.9 podemos observar a duração média de cada pergunta relativamente à posição dos modelos nos testes e à dificuldade do modelo.

Tabela 8.9: Estatísticas dos tempos das perguntas do grupo B

	Duração Média	Duração Mínima	Duração Máxima
Q1	00:37min	00:07min	02:27min
Q2	00:33min	00:01min	02:18min
Q3	00:24min	00:04min	02:42min
Q4	00:20min	00:02min	01:13min
Q5	00:23min	00:04min	01:17min

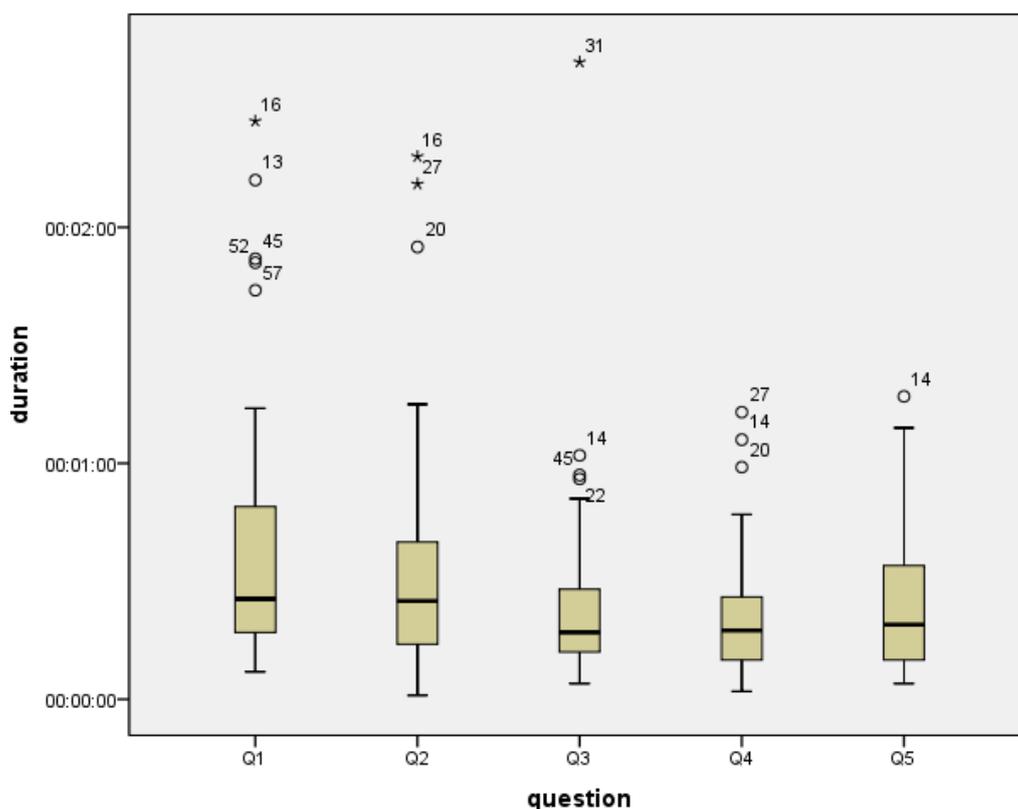


Figura 8.7: BoxPlot dos tempos dos participantes por pergunta do grupo B

Tabela 8.10: Estatísticas dos tempos das perguntas dos modelos do grupo B

	Duração Média	Duração Mínima	Duração Máxima
modelo	Q1 00:30min	00:07min	01:52min
fácil	Q2 00:17min	00:01min	00:58min
	Q3 00:17min	00:04min	01:02min
	Q4 00:15min	00:02min	01:06min

Continuação na próxima página

Tabela 8.10 – Continuação da página anterior

	Q5	00:13min	00:04min	01:17min
modelo	Q1	00:37min	00:07min	02:12min
médio	Q2	00:32min	00:11min	01:15min
	Q3	00:30min	00:09min	02:42min
	Q4	00:24min	00:07min	01:13min
	Q5	00:28min	00:09min	01:09min
modelo	Q1	00:44min	00:13min	02:27min
difícil	Q2	00:49min	00:11min	02:07min
	Q3	00:24min	00:06min	00:46min
	Q4	00:22min	00:05min	00:41min
	Q5	00:29min	00:06min	00:55min

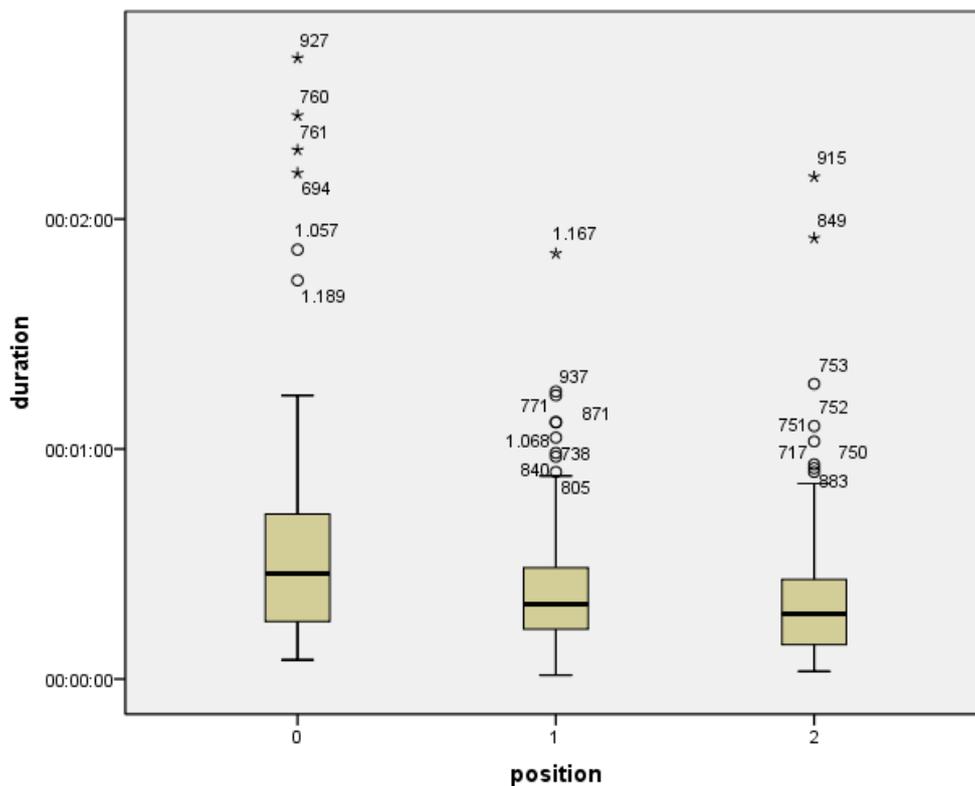


Figura 8.8: BoxPlot dos tempos dos participantes por posição do grupo B

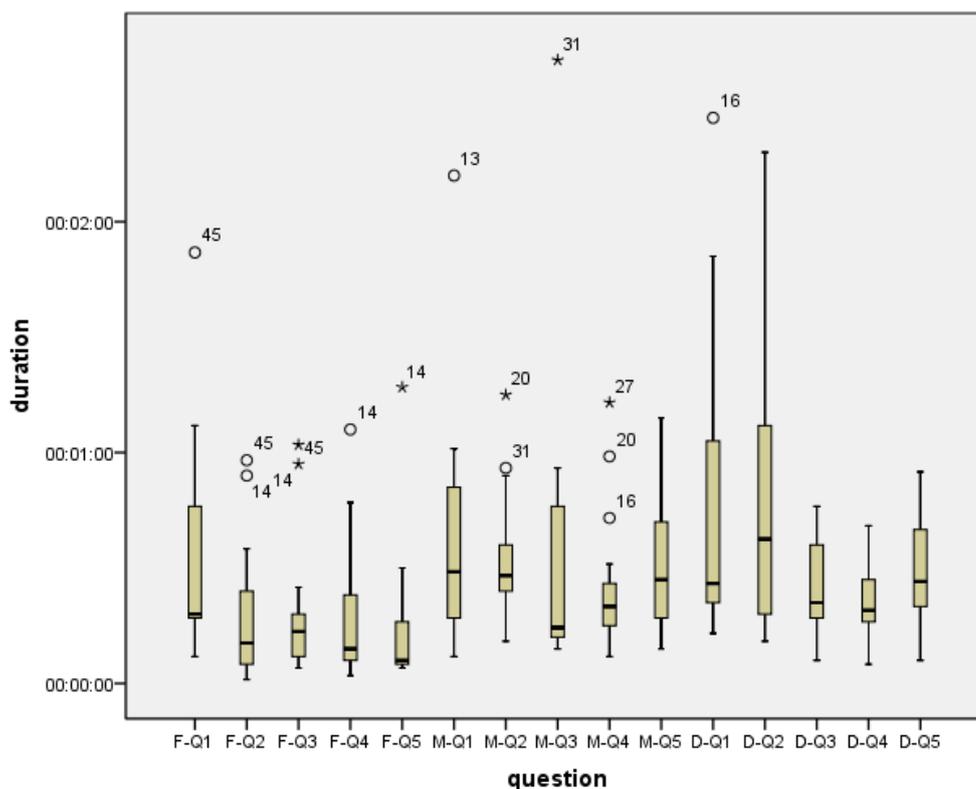


Figura 8.9: BoxPlot dos tempos dos participantes por dificuldade do grupo B

Grupo C Na Tabela 8.11 e na Figura 8.10 é possível ver dados relativos aos tempos de realização dos testes do grupo C por parte dos participantes, sendo que os testes são separados pela posição em que é apresentado cada um dos três modelos dos testes.

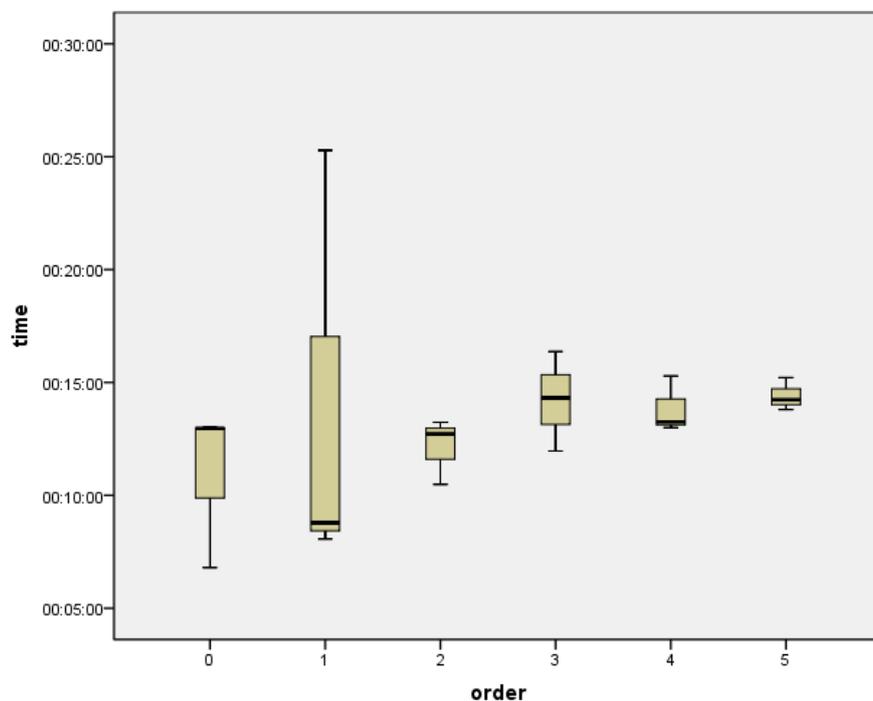
Tabela 8.11: Estatísticas descritivas dos tempos do grupo C

Ordem	Duração Média	Modelo	Duração Média	Duração Mínima	Duração Máxima
O1	10:55min	Fácil	02:03min	01:51min	02:25min
		Médio	04:01min	02:15min	04:20min
		Difícil	03:25min	01:50min	04:50min
O2	14:02min	Fácil	03:30min	02:09min	05:53min
		Médio	04:08min	02:07min	07:56min
		Difícil	04:59min	02:07min	10:04min
O3	12:08min	Fácil	01:38min	01:07min	02:23min
		Médio	04:14min	03:43min	04:54min
		Difícil	04:51min	03:56min	06:07min
O4	14:13min	Fácil	01:58min	01:35min	02:25min
		Médio	05:30min	04:39min	06:55min

Continuação na próxima página

Tabela 8.11 – *Continuação da página anterior*

		Difícil	05:23min	03:57min	06:07min
		Fácil	01:49min	01:18min	02:13min
O5	13:50min	Médio	03:29min	03:21min	03:46min
		Difícil	06:35min	05:35min	06:19min
		Fácil	01:35min	01:05min	02:02min
O6	14:25min	Médio	04:58min	04:49min	05:20min
		Difícil	06:16min	05:07min	07:32min

Figura 8.10: *BoxPlot* dos tempos dos participantes por ordem do grupo C

O grupo C em média demora cerca de treze minutos e quinze segundos a ser concluído. Este é o teste com uma maior média de duração dos testes, sendo que deste tempo em média o utilizador despende mais de dois terços do seu tempo entre o modelo médio e difícil, verificando-se que não existe uma grande discrepância entre o tempo despendido para o modelo médio e difícil como era expectável. De um modo geral entre as posições de análise dos modelos, não existe uma mudança drástica dos tempos, tal como é possível verificar no grupo A.

Para além da duração do grupo C, temos ainda o tempo de resposta às cinco perguntas de cada modelo. Nas seguintes tabelas 8.12 e 8.11 e gráficos 8.11, 8.12 e 8.13 podemos observar a duração média de cada pergunta relativamente à posição dos modelos nos testes e à dificuldade do modelo.

Tabela 8.12: Estatísticas dos tempos das perguntas do grupo C

	Duração Média	Duração Mínima	Duração Máxima
Q1	01:03min	00:09min	05:39min
Q2	00:41min	00:11min	01:41min
Q3	00:52min	00:09min	02:48min
Q4	00:32min	00:04min	02:32min
Q5	00:54min	00:09min	02:28min

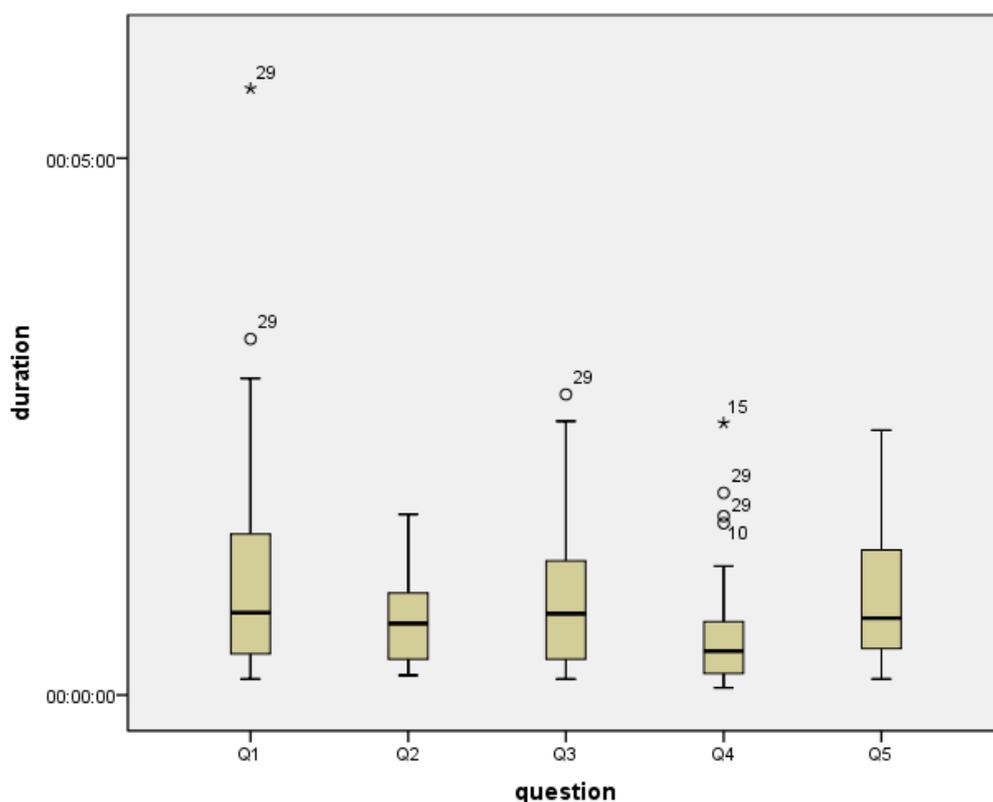


Figura 8.11: BoxPlot dos tempos dos participantes por pergunta do grupo C

Tabela 8.13: Estatísticas dos tempos das perguntas dos modelos do grupo C

		Duração Média	Duração Mínima	Duração Máxima
modelo	Q1	00:50min	00:11min	03:19min
fácil	Q2	00:19min	00:11min	00:35min
	Q3	00:22min	00:09min	00:56min

Continuação na próxima página

Tabela 8.13 – Continuação da página anterior

	Q4	00:15min	00:04min	00:37min
	Q5	00:29min	00:09min	02:23min
modelo	Q1	01:06min	00:09min	02:57min
médio	Q2	00:50min	00:18min	01:41min
	Q3	01:01min	00:17min	02:17min
	Q4	00:35min	00:07min	01:53min
	Q5	00:56min	00:23min	01:56min
modelo	Q1	01:14min	00:10min	05:39min
difícil	Q2	00:54min	00:28min	01:32min
	Q3	01:12min	00:11min	02:48min
	Q4	00:46min	00:09min	02:23min
	Q5	01:18min	00:29min	02:28min

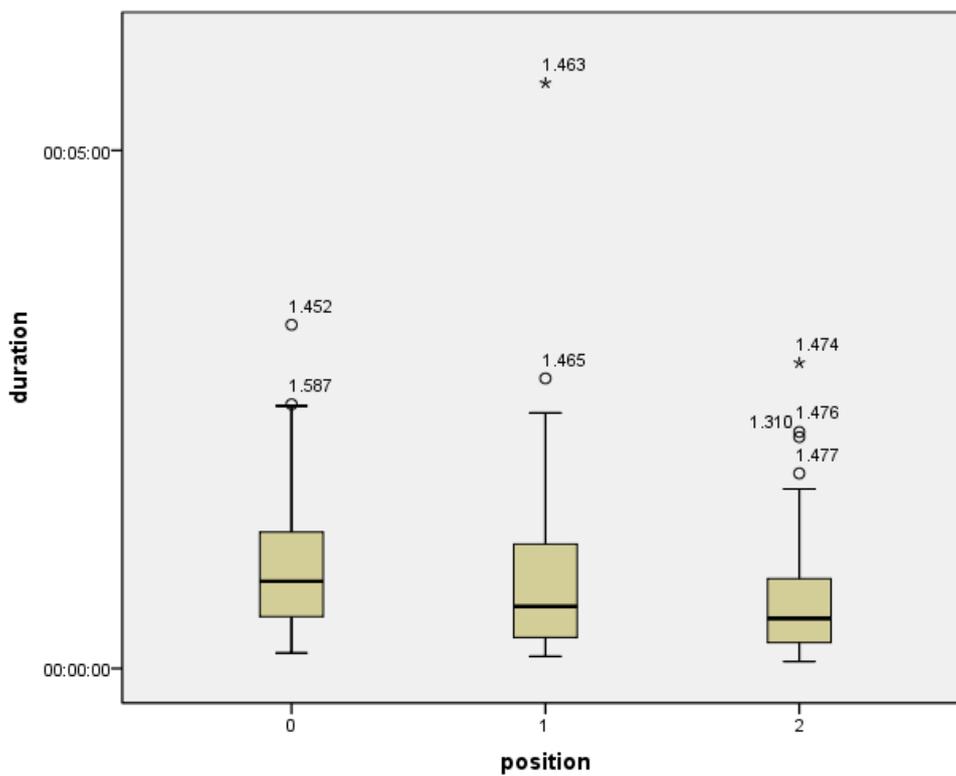


Figura 8.12: BoxPlot dos tempos dos participantes por posição do grupo C

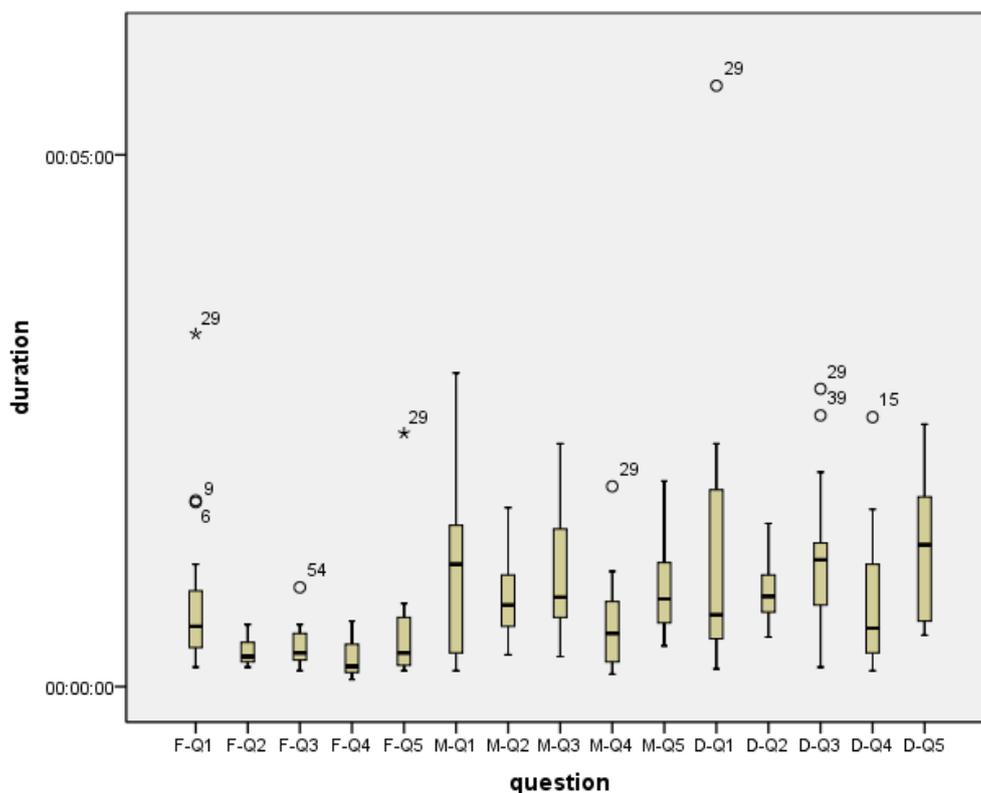


Figura 8.13: *BoxPlot* dos tempos dos participantes por dificuldade do grupo C

Análise

De modo a avaliar a normalidade dos dados referentes ao tempo de resposta às perguntas dos modelos dos testes realizados, procedeu-se à realização do teste de Shapiro-Wilk. Pode-se observar que não existe uma normalidade nos dados em todas as perguntas dos testes com $p < 0.05$.

Na [Tabela 8.14](#) é apresentada a análise descritiva da variável duração nos diferentes grupos de teste. E nas nas tabelas [8.15](#) e [8.16](#) é apresenta a análise estatística da variável duração, mas desta vez é diferenciada tendo em conta a posição ou a dificuldade dos modelos dos testes.

Tabela 8.14: Estatísticas descritivas dos tempos dos diferentes grupos

	N	Mínimo	Máximo	Média	Desvio	Variância	Skewness	Kurtosis
A	284	00:02min	04:25min	00:54min	00:49min	2464,493	1,937	4,130
B	270	00:01min	02:42min	00:27min	00:25min	638,817	2,434	7,710
C	270	00:04min	05:39min	00:48min	00:40min	1644,926	2,273	9,901

Posição

Tabela 8.15: Estatísticas descritivas dos tempos dos diferentes grupos por posição

	P	N	Mínimo	Máximo	Média	Desvio	Variância	Skewness	Kurtosis
A	1º	95	00:02min	04:04min	01:07min	00:51min	2605,180	1,363	1,792
A	2º	94	00:03min	04:18min	00:49min	00:46min	2160,115	2,304	6,463
A	3º	95	00:04min	04:25min	00:45min	00:49min	2417,502	2,536	7,229
B	1º	90	00:05min	02:42min	00:36min	00:31min	967,510	2,201	5,444
B	2º	90	00:01min	01:51min	00:24min	00:19min	366,565	1,837	4,446
B	3º	90	00:02min	02:11min	00:22min	00:22min	491,165	2,537	8,351
C	1º	90	00:09min	03:19min	00:59min	00:38min	1471,728	1,192	1,3674
C	2º	90	00:07min	05:39min	00:49min	00:49min	2214,460	3,069	15,344
C	3º	90	00:04min	02:57min	00:38min	00:32min	1061,268	1,843	4,016

De modo a avaliar a normalidade dos dados referentes ao tempo de resposta das perguntas dos grupos de teste em relação à posição dos modelos, procedeu-se à realização do teste de Shapiro-Wilk. Caso exista uma normalidade nos dados, pode-se proceder a uma análise de qui quadrado e o teste *oneway* entre grupos anova, de modo a perceber o efeito significativo dos dados.

Pode-se observar que não existe uma normalidade nos dados do grupo A nas perguntas Q1, Q2 e Q5 em todas as posições dos testes, na pergunta Q3 na segunda e terceira posição e na pergunta Q4 na terceira posição, com $p < 0.05$. Por outro lado, existe normalidade nos dados do grupo A no modelo na primeira posição na pergunta Q3 e na primeira e segunda posição da pergunta Q4, com $p > 0.05$.

Calculamos um teste de independência do qui quadrado comparando o efeito da ordem da execução dos modelos do teste no que toca ao tempo de resposta da pergunta Q3 do grupo A. Não foi encontrada uma relação significativa ($X^2(94) = 92,000$, $p = 0.539$) entre a dificuldade dos modelos e o tempo de resposta à pergunta.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q3 com a posição dos modelos do grupo A. Não existe um efeito significativo entre a dificuldade dos modelos no teste e o tempo de resposta à pergunta Q3 com $p < 0.05$ nas três condições ($F(2,54) = 1.388$, $p = 0.258$).

Calculamos um teste de independência do qui quadrado comparando o efeito da ordem da execução dos modelos do teste no que toca ao tempo de resposta da pergunta Q4 do grupo A. Não foi encontrada uma relação significativa ($X^2(90) = 90,000$, $p = 0.480$) entre a dificuldade dos modelos e o tempo de resposta à pergunta.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q4 com a posição dos modelos do grupo A. Não existe um efeito significativo entre a dificuldade dos modelos no teste e o tempo de resposta à pergunta Q4 com $p < 0.05$ nas três condições ($F(2,54) = 2.897$, $p = 0.064$).

Em suma, não foram encontrados efeitos significativos no grupo A no que toca à duração de resposta às perguntas dos testes e à posição dos modelos no teste.

Pode-se observar que não existe uma normalidade nos dados do grupo B nas perguntas Q1 em todas as posições dos testes, na pergunta Q2 na primeira e terceira posição, na pergunta Q3 na primeira e terceira posição, na pergunta Q4 na segunda e terceira posição e na pergunta Q5 na terceira posição, com $p < 0.05$. Por outro lado, existe normalidade nos dados do grupo B no modelo na segunda posição na pergunta Q2 e Q3, na primeira posição da pergunta Q4 e na primeira e segunda posição da pergunta Q5, com $p > 0.05$.

Calculamos um teste de independência do qui quadrado comparando o efeito da ordem da execução dos modelos do teste no que toca ao tempo de resposta da pergunta Q2 do grupo B. Não foi encontrada uma relação significativa ($X^2(76) = 73,000$, $p = 0.576$) entre a dificuldade dos modelos e o tempo de resposta à pergunta.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q2 com a posição dos modelos do grupo B. Não existe um efeito significativo entre a dificuldade dos modelos no teste e o tempo de resposta à pergunta Q2 com $p < 0.05$ nas três condições ($F(2,51) = 0.315$, $p = 0.732$).

Calculamos um teste de independência do qui quadrado comparando o efeito da ordem da execução dos modelos do teste no que toca ao tempo de resposta da pergunta Q3 do grupo B. Não foi encontrada uma relação significativa ($X^2(58) = 71,600$, $p = 0,108$) entre a dificuldade dos modelos e o tempo de resposta à pergunta.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q3 com a posição dos modelos do grupo B. Não existe um efeito significativo entre a dificuldade dos modelos no teste e o tempo de resposta à pergunta Q3 com $p < 0.05$ nas três condições ($F(2,51) = 0.194$, $p = 0.078$).

Calculamos um teste de independência do qui quadrado comparando o efeito da ordem da execução dos modelos do teste no que toca ao tempo de resposta da pergunta Q4 do grupo B. Não foi encontrada uma relação significativa ($X^2(64) = 67,000$, $p = 0.375$) entre a dificuldade dos modelos e o tempo de resposta à pergunta.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q4 com a posição dos modelos do grupo B. Não existe um efeito significativo entre a dificuldade dos modelos no teste e o tempo de resposta à pergunta Q4 com $p < 0.05$ nas três condições ($F(2,51) = 1.311$, $p = 0.283$).

Calculamos um teste de independência do qui quadrado comparando o efeito da ordem da execução dos modelos do teste no que toca ao tempo de resposta da pergunta Q5 do grupo B. Não foi encontrada uma relação significativa ($X^2(64) = 69,900$, $p = 0.286$) entre a dificuldade dos modelos e o tempo de resposta à pergunta.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q5 com a posição dos modelos do grupo B. Não existe um efeito significativo entre a dificuldade dos modelos no teste e o tempo de resposta à pergunta Q5 com $p < 0.05$ nas três condições ($F(2,51) = 0.291$, $p = 0.364$).

Em suma, não foram encontrados efeitos significativos no grupo B no que toca à duração de resposta às perguntas dos testes e à posição dos modelos no teste.

Pode-se observar que não existe uma normalidade nos dados do grupo C nas perguntas Q1 e Q4 em todas as posições dos testes, na pergunta Q2 na terceira posição, na pergunta Q3 na segunda e terceira posição, na pergunta Q5 na segunda posição, com $p < 0.05$. Por outro lado, existe normalidade nos dados do grupo C no modelo na primeira e segunda posição na pergunta Q1, na primeira posição da pergunta Q3 e na primeira e terceira posição da pergunta Q5, com $p > 0.05$.

Calculamos um teste de independência do qui quadrado comparando o efeito da ordem da execução dos modelos do teste no que toca ao tempo de resposta da pergunta Q1 do grupo C. Não foi encontrada uma relação significativa ($X^2(90) = 93,000$, $p = 0.393$) entre a dificuldade dos modelos e o tempo de resposta à pergunta.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q1 com a posição dos modelos do grupo C. Não existe um efeito significativo entre a dificuldade dos modelos no teste e o tempo de resposta à pergunta Q1 com $p < 0.05$ nas três condições ($F(2,51) = 0.553$, $p = 0.501$).

Calculamos um teste de independência do qui quadrado comparando o efeito da ordem da execução dos modelos do teste no que toca ao tempo de resposta da pergunta Q3 do grupo C. Não foi encontrada uma relação significativa ($X^2(86) = 86,000$, $p = 0.480$) entre a dificuldade dos modelos e o tempo de resposta à pergunta.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q3 com a posição dos modelos do grupo C. Não existe um efeito significativo entre a dificuldade dos modelos no teste e o tempo de resposta à pergunta Q3 com $p < 0.05$ nas três condições ($F(2,51) = 0.205$, $p = 0.160$).

Calculamos um teste de independência do qui quadrado comparando o efeito da ordem da execução dos modelos do teste no que toca ao tempo de resposta da pergunta Q5 do grupo C. Não foi encontrada uma relação significativa ($X^2(82) = 90,000$, $p = 0.256$) entre a dificuldade dos modelos e o tempo de resposta à pergunta.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q5 com a posição dos modelos do grupo C. Não existe um efeito significativo entre a dificuldade dos modelos no teste e o tempo de resposta à pergunta Q5 com $p < 0.05$ nas três condições ($F(2,51) = 3.650$, $p = 0.033$). Apesar do teste *oneway* demonstrar que existe uma relação significativa, realizou-se o teste efeito do tamanho ($R = 0.125$) e concluiu-se que as diferenças encontradas não são muito relevantes para o teste em causa.

Em suma, não foram encontrados efeitos significativos no grupo C no que toca à duração de resposta às perguntas dos testes e à posição dos modelos no teste.

Dificuldade

Tabela 8.16: Estatísticas descritivas dos tempos dos diferentes grupos por dificuldade

	P	N	Mínimo	Máximo	Média	Desvio	Variância	Skewness	Kurtosis
A	fácil	95	00:03min	03:25min	00:40min	00:39min	1572,212	2,059	5,195
A	méd	94	00:02min	04:18min	00:51min	00:39min	1555,413	2,458	10,180
A	difíc	95	00:06min	04:25min	01:10min	01:02min	3845,359	1,389	1,113
B	fácil	90	00:01min	01:52min	00:19min	00:19min	392,977	2,182	5,534
B	médi	90	00:07min	02:42min	00:30min	00:24min	593,858	2,775	11,258
B	difíc	90	00:05min	02:27min	00:34min	00:28min	817,628	2,317	5,573
C	fácil	90	00:04min	03:19min	00:27min	00:28min	802,028	3,792	17,862
C	méd	90	00:07min	02:57min	00:54min	00:33min	1126,460	0,943	0,931
C	difíc	90	00:09min	05:39min	01:05min	00:47min	2287,268	2,468	11,030

Pode-se observar que não existe uma normalidade nos dados do grupo A nas perguntas Q2, Q3 e Q5 em todas as posições dos testes, na pergunta Q1 na primeira e terceira posição e na pergunta Q4 na terceira posição, com $p < 0.05$. Por outro lado, existe normalidade nos dados do grupo A no modelo na segunda posição na pergunta Q1 e na primeira e segunda posição da pergunta Q4, com $p > 0.05$.

Calculamos um teste de independência do qui quadrado comparando o efeito da dificuldade dos modelos do teste no que toca ao tempo de resposta da pergunta Q1 do grupo A. Não foi encontrada uma relação significativa ($X^2(2) = 92,000$, $p = 0.364$) entre a dificuldade dos modelos e a duração.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q1 com a dificuldade dos modelos do grupo A. Não existe um efeito significativo entre as posições dos modelos no teste e o tempo de resposta à pergunta Q1 com $p < 0.05$ nas três condições ($F(2,54) = 2,87$, $p = 0.05$).

Calculamos um teste de independência do qui quadrado comparando o efeito da dificuldade dos modelos do teste no que toca ao tempo de resposta da pergunta Q4 do grupo A. Não foi encontrada uma relação significativa ($X^2(2) = 90,000$, $p = 0.480$) entre a dificuldade dos modelos e a duração.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q4 com a dificuldade dos modelos do grupo A. Não existe um efeito significativo entre as posições dos modelos no teste e o tempo de resposta à pergunta Q4 com $p < 0.05$ nas três condições ($F(2,54) = 0,414$, $p = 0.065$).

Em suma, não foram encontrados efeitos significativos no grupo A no que toca à duração de resposta às perguntas dos testes e à dificuldade dos modelos no teste.

Pode-se observar que não existe uma normalidade nos dados do grupo B nas perguntas Q1 e Q2 em todas as posições dos testes, na pergunta Q3 na primeira e segunda posição, na pergunta Q4 na primeira e segunda posição e na pergunta Q5 na primeira posição,

com $p < 0.05$. Por outro lado, existe normalidade nos dados do grupo B no modelo na terceira posição na pergunta Q3, na terceira posição da pergunta Q4 e na segunda e terceira posição da pergunta Q5, com $p > 0.05$.

Calculamos um teste de independência do qui quadrado comparando o efeito da dificuldade dos modelos do teste no que toca ao tempo de resposta da pergunta Q3 do grupo B. Não foi encontrada uma relação significativa ($X^2(58) = 63,200$, $p = 0.298$) entre a dificuldade dos modelos e a duração.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q3 com a dificuldade dos modelos do grupo B. Não existe um efeito significativo entre as posições dos modelos no teste e o tempo de resposta à pergunta Q3 com $p < 0.05$ nas três condições ($F(2,51) = 1.232$, $p = 0.300$).

Calculamos um teste de independência do qui quadrado comparando o efeito da dificuldade dos modelos do teste no que toca ao tempo de resposta da pergunta Q4 do grupo B. Não foi encontrada uma relação significativa ($X^2(64) = 75,200$, $p = 0.164$) entre a dificuldade dos modelos e a duração.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q4 com a dificuldade dos modelos do grupo B. Não existe um efeito significativo entre as posições dos modelos no teste e o tempo de resposta à pergunta Q4 com $p < 0.05$ nas três condições ($F(2,51) = 1.600$, $p = 0.212$).

Calculamos um teste de independência do qui quadrado comparando o efeito da dificuldade dos modelos do teste no que toca ao tempo de resposta da pergunta Q5 do grupo B. Não foi encontrada uma relação significativa ($X^2(64) = 77,200$, $p = 0.124$) entre a dificuldade dos modelos e a duração.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q5 com a dificuldade dos modelos do grupo B. Existe um efeito significativo entre as posições dos modelos no teste e o tempo de resposta à pergunta Q5 com $p < 0.05$ nas três condições ($F(2,51) = 5,877$, $p = 0.005$). Apesar do teste *oneway* demonstrar que existe uma relação significativa, realizou-se o teste efeito do tamanho ($R = 0.187$) e concluiu-se que as diferenças encontradas não são muito relevantes para o teste em causa.

Em suma, não foram encontrados efeitos significativos no grupo B no que toca à duração de resposta às perguntas dos testes e a dificuldade dos modelos no teste.

Pode-se observar que não existe uma normalidade nos dados do grupo C na pergunta Q1 na primeira e terceira posição, na pergunta Q3 na primeira posição, na pergunta Q4 na segunda e terceira posição e na pergunta Q5 na primeira e segunda posição, com $p < 0.05$. Por outro lado, existe normalidade nos dados do grupo C no modelo na segunda posição na pergunta Q1, em todas as posições da pergunta Q2, na segunda e terceira posição da pergunta Q3 e na primeira posição da pergunta Q4 e na terceira posição da questão Q5, com $p > 0.05$.

Calculamos um teste de independência do qui quadrado comparando o efeito da dificuldade dos modelos do teste no que toca ao tempo de resposta da pergunta Q1 do grupo C. Não foi encontrada uma relação significativa ($X^2(90) = 99,000$, $p = 0.242$) entre a dificuldade dos modelos e a duração.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q1 com a dificuldade dos modelos do grupo C. Existe um efeito significativo entre as posições dos modelos no teste e o tempo de resposta à pergunta Q1 com $p < 0.05$ nas três condições ($F(2,51) = 0,772$, $p = 468$).

Calculamos um teste de independência do qui quadrado comparando o efeito da dificuldade dos modelos do teste no que toca ao tempo de resposta da pergunta Q2 do grupo C. Não foi encontrada uma relação significativa ($X^2(90) = 99,000$, $p = 0.242$) entre a dificuldade dos modelos e a duração.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q2 com a dificuldade dos modelos do grupo C. Existe um efeito significativo entre as posições dos modelos no teste e o tempo de resposta à pergunta Q2 com $p < 0.05$ nas três condições ($F(2,51) = 23,004$, $p = 0.000$). Apesar do teste *oneway* demonstrar que existe uma relação significativa, realizou-se o teste efeito do tamanho ($R = 0.474$) e concluiu-se que as diferenças encontradas são ligeiramente relevantes para o teste em causa.

Calculamos um teste de independência do qui quadrado comparando o efeito da dificuldade dos modelos do teste no que toca ao tempo de resposta da pergunta Q3 do grupo C. Não foi encontrada uma relação significativa ($X^2(86) = 95,000$, $p = 0.238$) entre a dificuldade dos modelos e a duração.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q3 com a dificuldade dos modelos do grupo C. Existe um efeito significativo entre as posições dos modelos no teste e o tempo de resposta à pergunta Q3 com $p < 0.05$ nas três condições ($F(2,51) = 12,123$, $p = 0.000$). Apesar do teste *oneway* demonstrar que existe uma relação significativa, realizou-se o teste efeito do tamanho ($R = 0.322$) e concluiu-se que as diferenças encontradas são ligeiramente relevantes para o teste em causa.

Calculamos um teste de independência do qui quadrado comparando o efeito da dificuldade dos modelos do teste no que toca ao tempo de resposta da pergunta Q4 do grupo C. Não foi encontrada uma relação significativa ($X^2(66) = 76,500$, $p = 0.177$) entre a dificuldade dos modelos e a duração.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q4 com a dificuldade dos modelos do grupo C. Existe um efeito significativo entre as posições dos modelos no teste e o tempo de resposta à pergunta Q4 com $p < 0.05$ nas três condições ($F(2,51) = 5,796$, $p = 0.005$). Apesar do teste *oneway* demonstrar que existe uma relação significativa, realizou-se o teste efeito do tamanho ($R = 0.185$) e concluiu-se que as diferenças encontradas não são relevantes para o teste em causa.

Calculamos um teste de independência do qui quadrado comparando o efeito da dificuldade dos modelos do teste no que toca ao tempo de resposta da pergunta Q5 do grupo C. Não foi encontrada uma relação significativa ($X^2(66) = 76,500$, $p = 0.177$) entre a dificuldade dos modelos e a duração.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q5 com a dificuldade dos modelos do grupo C. Existe um efeito significativo entre as posições dos modelos no teste e o tempo de resposta à pergunta Q5 com $p < 0.05$ nas três condições ($F(2,51) = 9,980$, $p = 0.000$). Apesar do teste *oneway* demonstrar que existe uma relação significativa, realizou-se o teste efeito do tamanho ($R = 0.281$) e concluiu-se que as diferenças encontradas não são relevantes para o teste em causa.

Em suma, foi encontrado um efeito significativo na pergunta Q2 e Q3 no grupo C no que toca à duração de resposta às perguntas dos testes e à dificuldade dos modelos no teste. Apesar de nas perguntas Q4 e Q5 o efeito significativo não ser relevante, os modelos do grupo C provocam um efeito no tempo de resposta às perguntas.

8.2.2 Correção

Nesta secção serão apresentados os dados dos testes relativamente às resposta corretas das diferentes perguntas dos diferentes grupos e as suas respetivas análises estatísticas quando aplicáveis.

Dados

Grupo A Na Tabela 8.17 e e Figura 8.14 é possível ver dados relativos às respostas dadas pelos participantes nos seus testes. As respostas são apresentadas também recorrendo à mesma divisão que a tabela anterior, para que possa ser possível retirar ilações sobre o impacto da dificuldade dos modelos.

Tabela 8.17: Percentagem de respostas corretas do grupo A por pergunta

Modelo	Posição	Q1	Q2	Q3	Q4	Q5
Fácil	1	83,3%	100%	83,3%	0%	50%
Fácil	2	100%	83,3%	16,67%	66,67%	83,3%
Fácil	3	71,43%	71,43%	28,57%	71,43%	71,43%
Médio	1	54,14%	54,14%	14,29%	0%	42,86%
Médio	2	33,3%	66,67%	16,67%	0%	100%
Médio	3	83,3%	66,67%	33,3%	16,67%	66,67%
Difícil	1	33,3%	66,67%	16,67%	33,3%	66,67%
Difícil	2	87,71%	71,43%	28,57%	57,14%	42,86%
Difícil	3	83,3%	33,3%	50%	33,3%	66,67%

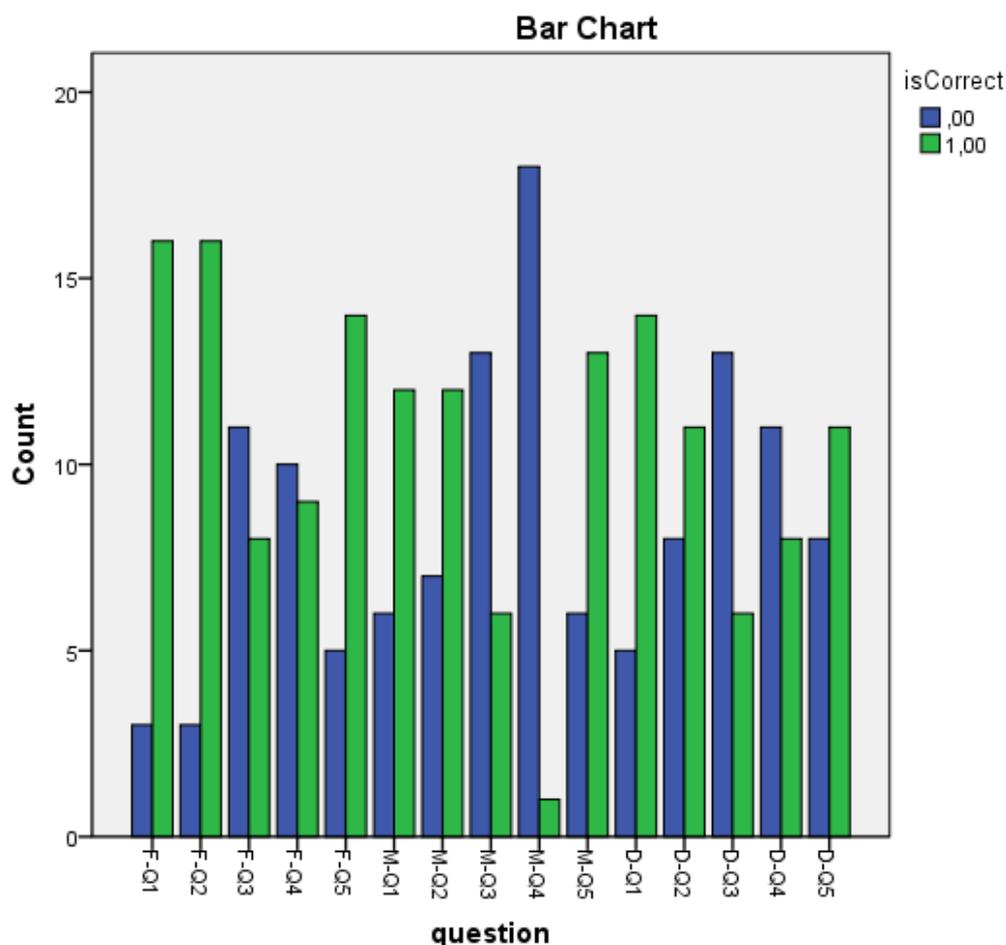


Figura 8.14: BarChart das respostas dos participantes do grupo A

Olhando agora para a percentagem de respostas corretas dos vários modelos e para os tempos de conclusão dos modelos, podemos concluir que a velocidade de conclusão dos modelos não significa que as respostas dadas pelo participante estejam corretas. Em relação à primeira pergunta Q1 - *Quantos participantes existem no modelo?* é possível observar que independentemente do modelo não existe uma melhoria significativa de respostas corretas ao decorrer dos modelos. Quanto à segunda pergunta Q2 - *Todos os participantes são essenciais para a conclusão do processo?* existe um decréscimo de respostas corretas à medida que se avança na dificuldade dos modelos. Quando os testes começam pelos modelos mais fáceis, não é perceptível ao utilizador quais são os participantes que obrigatoriamente executam o processo em causa.

No que diz respeito à terceira pergunta, Q3 - *Quantas atividades são ativadas através de um acontecimento externo ou evento?* esta possui uma melhor média de respostas corretas quando o modelo apresentado é o modelo fácil, ou seja, quando o nível de dificuldade aumenta o participante não tem noção do que representa um acontecimento externo ou evento. Na quarta pergunta Q4 - *Quantos participantes utilizam/geram dados para as*

suas atividades processarem/utilizarem? existe uma melhoria significativa com a subida dos níveis de dificuldade. Por último, a quinta pergunta Q5 - *Todos os fluxos de sequência divididos voltam a juntar-se?* a taxa de respostas corretas mantém-se equivalente ao longo dos modelos.

Grupo B na Tabela 8.18 e Figura 8.15 é possível ver dados relativos às respostas dadas pelos participantes nos seus testes. As respostas estão apresentadas recorrendo à mesma divisão que a tabela anterior, para que possa ser possível retirar ilações sobre o impacto da dificuldade dos modelos.

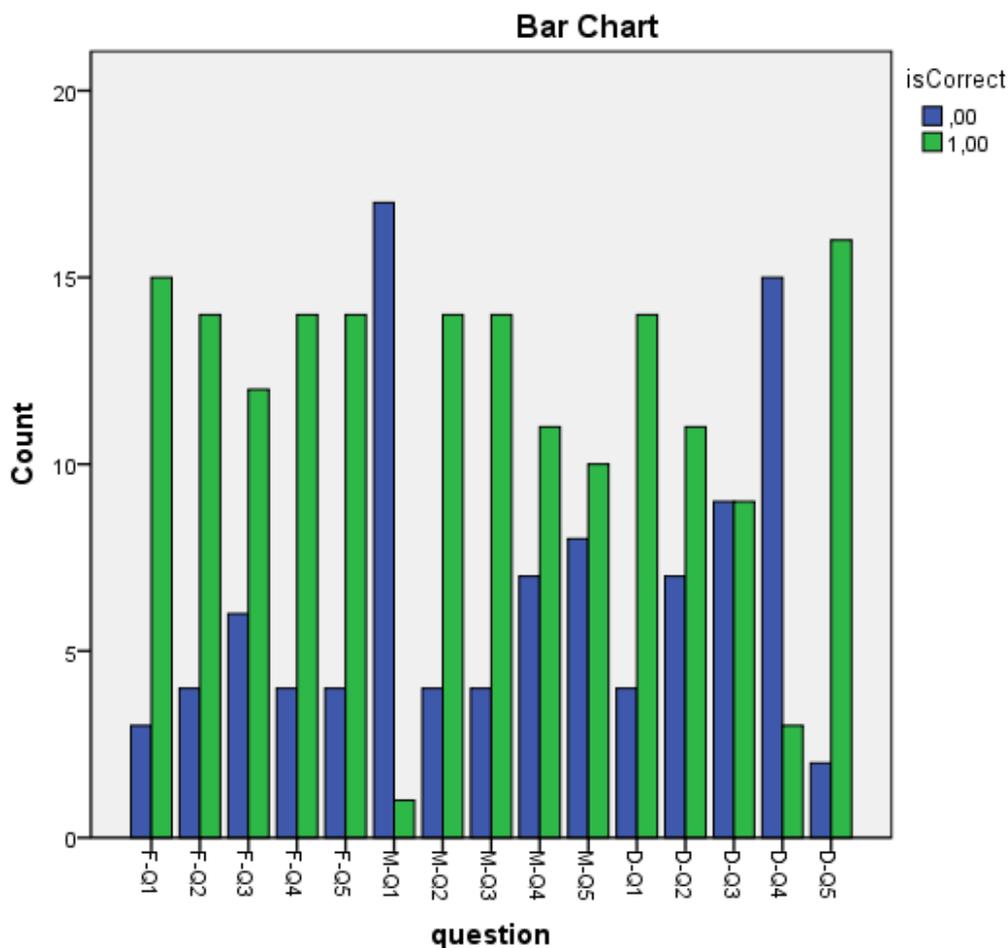


Figura 8.15: *BarChart* das respostas dos participantes do grupo B

Tabela 8.18: Percentagem de respostas corretas do grupo B por pergunta

Modelo	Posição	Q1	Q2	Q3	Q4	Q5
Fácil	1	83,3%	83,3%	66,67%	83,3%	83,3%
Fácil	2	66,67%	50%	50%	50%	50%
Fácil	3	100%	100%	83,3%	100%	100%

Continuação na próxima página

Tabela 8.18 – *Continuação da página anterior*

Médio	1	16,67%	66,67%	50%	33,3%	50%
Médio	2	0%	83,3%	83,3%	83,3%	50%
Médio	3	0%	83,3%	100%	67,67%	67,67%
Difícil	1	83,3%	66,67%	66,67%	33,3%	83,3%
Difícil	2	83,3%	66,67%	67,67%	33,3%	100%
Difícil	3	83,3%	67,67%	33,3%	16,67%	100%

No que diz respeito à relação entre as respostas corretas e o tempo de conclusão dos testes e dos modelos em particular, podemos observar que existem melhorias nas respostas dos modelos fácil quando este é visualizado numa posição mais avançada no teste, sendo que os tempos de resposta diminuí. Podemos então afirmar que quanto maior é a interação com os modelos, melhor é o tempo de resposta sendo estas feitas de um modo mais assertivo.

O modelo médio à semelhança do modelo fácil também tem um aumento de respostas corretas quando este é visualizado mais tarde nos testes. Por outro lado, o modelo difícil não possui grandes diferenças nos seus tempos de respostas nas suas diferentes posições nos testes, sendo que a taxa de respostas corretas é constante. Podemos afirmar então que, quando o modelo é teoricamente mais difícil, não basta existir unicamente uma interação rápida com os modelos, mas sim conhecer previamente as nomenclaturas para tornar a leitura mais rápida e eficiente.

Em relação às perguntas de um modo detalhado, a primeira pergunta *Q1 - Quantos sub-processos existem no processo?* diz-nos que os participantes não conseguem perceber ou identificar os sub-processos expandidos nos modelos, mas que os sub-processos sim são fáceis de identificar de um modo geral. Na segunda pergunta *Q2 - Algum dos sub-processos pode ser executado mais que uma vez? Se sim, quantos?*, as respostas corretas decrescem em relação à capacidade de identificar os sub-processos. No modelo médio existe o fenómeno de como o sub-processo não é repetido o utilizador acerta as respostas mesmo não sabendo identificar o sub-processo expandido. Na terceira pergunta, *Q3 - Todos os sub-processos são executados obrigatoriamente para a conclusão do processo?*, acontece o mesmo fenómeno que na segunda pergunta, o participante diminui ou mantém a taxa de respostas corretas.

Quanto à quarta pergunta, *Q4 - Algum dos sub-processos recorre a dados para ser executado? Se sim, quantos?*, consegue-se perceber que o utilizador não compreendeu o conceito e a taxa de respostas corretas diminui para esta pergunta em relação às restantes. Na quinta e última pergunta, *Q5 - Quantos sub processos são ativados através de um acontecimento externo ou evento?*, existe uma boa taxa de respostas correta no último a ser analisado, podemos então dizer que o utilizador começa a aprender com os modelos analisados.

Grupo C Na Tabela 8.19 e Figura 8.16 é possível ver dados relativos às respostas dadas pelos participantes nos seus testes. As respostas são apresentadas também recorrendo à mesma divisão que a tabela anterior, para que possa ser possível retirar ilações sobre o impacto da dificuldade dos modelos.

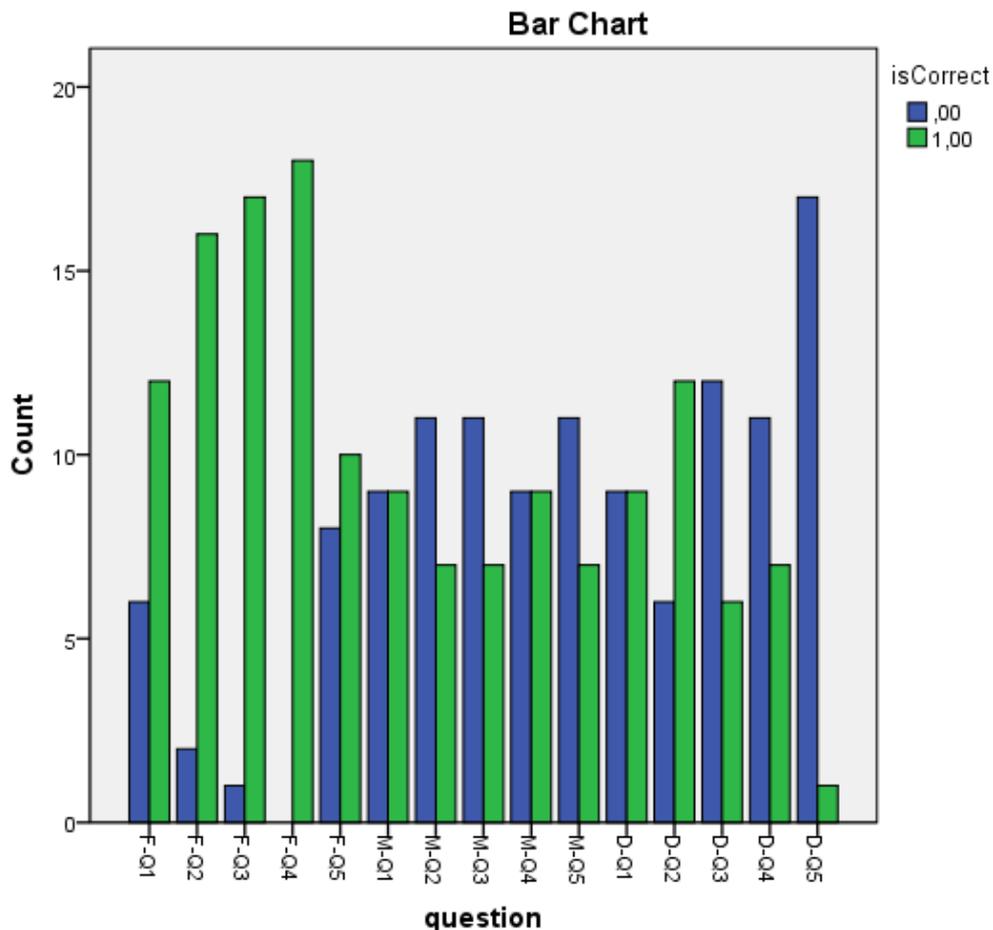


Figura 8.16: BarChart das respostas dos participantes do grupo C

Tabela 8.19: Percentagem de respostas corretas do grupo c por pergunta

Modelo	Posição	Q1	Q2	Q3	Q4	Q5
Fácil	1	50%	100%	100%	100%	33,3%
Fácil	2	50%	83,3%	83,3%	100%	50%
Fácil	3	100%	83,3%	100%	100%	83,3%
Médio	1	50%	50%	0%	33,3%	33,3%
Médio	2	50%	33,3%	66,67%	50%	50%
Médio	3	50%	33,3%	50%	66,67%	33,3%
Difícil	1	33,3%	50%	33,3%	33,3%	0%
Difícil	2	50%	66,67%	33,3%	16,67%	0%
Difícil	3	66,67%	83,3%	16,67%	66,67%	0%

Em relação às respostas corretas e à duração dos testes e dos seus respetivos modelos, pode-se observar que o modelo fácil demora em média dois minutos a ser concluído, mas possui grandes taxas de aprovação. No que diz respeito ao modelo médio e difícil o mesmo já não é verificável, os modelos que levam mais tempo de análise, mas essa duração não se reflete num aumento da qualidade das respostas. À semelhança dos restantes grupos, o primeiro nível destes, quando analisado em primeiro lugar demora sempre mais tempo que quando analisados em segundo ou em último lugar. Este facto, verifica-se em todos os níveis de dificuldade.

Agora em relação às perguntas em si, a primeira pergunta, *Q1 - Quantos fluxos paralelos existem no modelo?*, é perceptível apenas para metade dos participantes dos testes, sendo que a taxa de sucesso desta aumenta nos últimos modelos analisados, devido ao efeito de aprendizagem que surge durante os testes. A segunda pergunta, *Q2 - Quantos ciclos existem no processo?*, contém uma taxa de aprovação média, ou seja, no modelo mais fácil é perceptível a resposta enquanto que no modelo médio é mais difícil por parte do utilizador. Apesar da terceira pergunta, *Q3 - Existe alguma atividade que seja executada mais que uma vez? Se sim, quantas?*, estar relacionada com a pergunta anterior, esta não apresenta as mesmas percentagens de aprovação, ou seja, quando o modelo é difícil a perceção da resposta torna-se mais difícil de obter mas quando o modelo é fácil a resposta é fácil de calcular.

A quarta pergunta, *Q4 - Os fluxos de sequência divididos são posteriormente agregados com recurso a um gateway?*, possui uma boa taxa de aprovação quando o modelo é fácil e com uma apresentação simples, o mesmo já não se verifica quando o modelo torna-se mais difícil e conseqüentemente mais confuso. Por último na quinta pergunta, *Q5 - Quantos dos fluxos de sequência são ativados através de um evento?*, é possível concluir que com modelos difíceis é mais complicado de calcular a resposta enquanto que nos modelos de nível inferior as respostas corretas aumentam ligeiramente.

Análise

Posição De modo a avaliar a normalidade dos dados referentes à correção das respostas dadas nos testes em relação à posição do modelo em causa no teste, procedeu-se à realização do teste de Shapiro-Wilk. Pode-se observar que não existe uma normalidade nos dados em todas as perguntas dos testes do grupo A com $p < 0.05$. Pode-se observar que não existe uma normalidade nos dados em todas as perguntas dos testes do grupo B com $p < 0.05$. Pode-se observar que não existe uma normalidade nos dados em todas as perguntas dos testes do grupo C com $p < 0.05$.

Como os dados não possuem uma distribuição normal, estes não podem ser analisados com os testes de dados paramétricos como o teste do qui-quadrado.

Dificuldade De modo a avaliar a normalidade dos dados referentes à correção das respostas dadas nos testes em relação à dificuldade do modelo em causa no teste, procedeu-se à realização do teste de Shapiro-Wilk. Pode-se observar que não existe uma normalidade nos dados em todas as perguntas dos testes do grupo A com $p < 0.05$. Pode-se observar que não existe uma normalidade nos dados em todas as perguntas dos testes do grupo B com $p < 0.05$. Pode-se observar que não existe uma normalidade nos dados em todas as perguntas dos testes do grupo C com $p < 0.05$.

Como os dados não possuem uma distribuição normal, estes não podem ser analisados com os testes de dados paramétricos como o teste do qui-quadrado.

8.2.3 Correção vs Duração

De modo a avaliar a normalidade dos dados referentes ao tempo de resposta às perguntas dos modelos dos testes realizados relativamente à correção das respostas dadas, procedeu-se à realização do teste de Shapiro-Wilk.

Pode-se observar que não existe uma normalidade nos dados em quase todas as perguntas dos testes com $p < 0.05$. As perguntas onde existe uma normalidade nos dados é na pergunta Q5 do grupo B, pergunta Q2, Q3 e Q4 do grupo C.

Calculamos um teste de independência do qui quadrado comparando o efeito do tempo de resposta à pergunta dos modelos do teste no que toca à correção de resposta da pergunta Q5 do grupo B. Não foi encontrada uma relação significativa ($X^2(32) = 10,114$, $p = 0.154$) entre a dificuldade dos modelos e a duração.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q5 com a correção das respostas à pergunta dos modelos do grupo B. Não existe um efeito significativo entre a duração e o tempo de resposta à pergunta Q5 com $p < 0.05$ nas três condições ($F(2,52) = 0,908$, $p = 0.345$).

Calculamos um teste de independência do qui quadrado comparando o efeito do tempo de resposta à pergunta dos modelos do teste no que toca à correção de resposta da pergunta Q2 do grupo C. Não foi encontrada uma relação significativa ($X^2(38) = 37,922$, $p = 0.473$) entre a dificuldade dos modelos e a duração.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q2 com a correção das respostas à pergunta dos modelos do grupo C. Não existe um efeito significativo entre a duração e o tempo de resposta à pergunta Q2 com $p < 0.05$ nas três condições ($F(2,52) = 0,312$, $p = 0.579$).

Calculamos um teste de independência do qui quadrado comparando o efeito do tempo de resposta à pergunta dos modelos do teste no que toca à correção de resposta da pergunta Q3 do grupo C. Não foi encontrada uma relação significativa ($X^2(43) = 45,794$, $p = 0,357$) entre a dificuldade dos modelos e a duração.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q3 com a correção das respostas à pergunta dos modelos

do grupo C. Não existe um efeito significativo entre a duração e o tempo de resposta à pergunta Q3 com $p < 0.05$ nas três condições ($F(2,52) = 1,460$, $p = 0,232$).

Calculamos um teste de independência do qui quadrado comparando o efeito do tempo de resposta à pergunta dos modelos do teste no que toca à correção de resposta da pergunta Q4 do grupo C. Não foi encontrada uma relação significativa ($X^2(33) = 42,410$, $p = 0,126$) entre a dificuldade dos modelos e a duração.

O teste *oneway* entre grupos anova foi realizado para proceder à comparação entre o tempo de resposta da pergunta Q4 com a correção das respostas à pergunta dos modelos do grupo C. Existe um efeito significativo entre a duração e o tempo de resposta à pergunta Q4 com $p < 0.05$ nas três condições ($F(2,52) = 7,194$, $p = 0,010$). Devido à grande variedade dos valores dos tempos de resposta, não foi possível calcular o valor do efeito do tamanho.

Em suma, foi encontrado um efeito significativo na pergunta Q4 grupo C no que toca à duração e à correção das resposta às perguntas dos testes. Apesar de tudo não existe uma grande expressividade de efeitos significativos em todos os grupos de teste.

8.2.4 NASA-TLX

As seguintes análises pretendem fornecer dados relativamente ao estado psicológico dos participantes nos testes e das suas opiniões relativamente aos testes por si realizados. As tabelas apresentadas são compostas por quatro colunas, na primeira e segunda coluna é apresentada a organização dos modelos e posições por que foram analisados e nas restantes colunas é apresentada a respetiva resposta com mais ocorrências por parte dos participantes. Caso o valor de frequência da resposta não seja o suficientemente expressiva serão apresentados os dois resultados mais frequentes.

De modo a facilitar a apresentação dos resultados dos testes do NASA-TLX, será utilizada a seguinte tabela que descreve as seguintes legendas:

Tabela 8.20: Legenda testes NASA-TLX

Abreviatura	Correspondência	Abreviatura	Correspondência
CT	Concordo Totalmente	MB	Muito Baixo
C	Concordo	B	Baixo
SO	Sem opinião	N	Normal
D	Discordo	E	Elevado
DT	Discordo Totalmente	ME	Muito Elevado

Grupo A No grupo A, para cada um dos modelos e para cada uma das posições possíveis existem seis participantes, com a exceção da existência de mais um participante com para a ordem quatro dos testes, ou seja, com o nível médio primeiro lugar, o nível difícil em

segundo lugar e no último lugar o nível fácil.

Tabela 8.21: A exigência mental para este teste foi a adequada?

Modelo	P	Moda	Freq.
Fácil	1	CT/C	3/3
Fácil	2	CT	4
Fácil	3	CT	5
Médio	1	C	4
Médio	2	C	4
Médio	3	CT	5
Difícil	1	CT/C	3/2
Difícil	2	C	4
Difícil	3	CT	5

Tabela 8.22: A exigência física para este teste foi a adequada?

Modelo	P	Moda	Freq.
Fácil	1	CT	6
Fácil	2	CT	4
Fácil	3	CT	6
Médio	1	CT	5
Médio	2	CT	6
Médio	3	CT	5
Difícil	1	CT	4
Difícil	2	CT	6
Difícil	3	CT	5

De um modo geral, todos os participantes independentemente da ordem pela qual os modelos apareceram, acharam que a exigência mental dos modelos foi adequada.

À semelhança da pergunta anterior, os participantes concordaram de um modo geral que a exigência física dos testes foi a correta. Nestes testes a exigência física correspondia à capacidade de fazer o mínimo de movimento possível com o corpo e a cabeça. Isto devido à sensibilidade de recolha de dados do *eye-tracker*.

Tabela 8.23: A exigência temporal para este teste foi a adequada?

Modelo	P	Moda	Freq.
Fácil	1	CT	5
Fácil	2	CT/SO	3/2
Fácil	3	CT	6
Médio	1	CT	6
Médio	2	CT	4
Médio	3	CT	4
Difícil	1	CT	3
Difícil	2	CT	7
Difícil	3	CT	4

Tabela 8.24: Como considera o seu esforço aplicado durante o teste?

Modelo	P	Moda	Freq.
Fácil	1	N	5
Fácil	2	MB/B	3/3
Fácil	3	B/N	3/2
Médio	1	E/N/B	3/2/2
Médio	2	N	4
Médio	3	N	5
Difícil	1	N	5
Difícil	2	N/B	3/2
Difícil	3	N/E	3/2

Na sua grande maioria todos os participantes concordaram que o tempo dos testes foi adequado à sua exigência e finalidade, com a exceção de um participante, que discordou com a exigência temporal do teste do modelo difícil em primeiro lugar.

De um modo geral o modelo fácil leva a que os participantes classifiquem o seu esforço como normal, baixo ou muito baixo em alguns casos. Com o modelo médio como primeiro modelo, os participantes classificam na sua grande maioria como normal o seu esforço, sendo que quando o modelo médio aparece em primeiro lugar o nível de esforço varia de elevado a baixo. Apesar do modelo difícil aparecer em último lugar em alguns casos, alguns participantes sentiram que o esforço para o perceber ainda foi elevado.

Tabela 8.25: Qual o seu nível de frustração durante o teste?

Modelo	P	Moda	Freq.
Fácil	1	B/N	2/2
Fácil	2	MB	4
Fácil	3	N/B	3/2
Médio	1	MB/N/E	2/2/2
Médio	2	B	3
Médio	3	MB/B	3/2
Difícil	1	N	4
Difícil	2	MB/N	2/2
Difícil	3	N/E	2/2

Tabela 8.26: Como classificaria a sua performance no teste?

Modelo	P	Moda	Freq.
Fácil	1	N	5
Fácil	2	N	4
Fácil	3	N	4
Médio	1	B/N	3/3
Médio	2	N/B	3/3
Médio	3	N	4
Difícil	1	B	5
Difícil	2	N/B	3/3
Difícil	3	N	4

Em geral os participantes apresentam um nível de frustração normal ou baixo no modelo fácil e sobe ligeiramente no modelo médio e difícil, havendo casos de participantes com a frustração muito elevada quando o modelo difícil é o primeiro modelo a ver ou quando o modelo difícil é o último a aparecer.

Quanto à perceção do desempenho os participantes de um modo geral acham que tiveram um desempenho normal no modelo fácil. Quanto ao modelo médio o desempenho passa a ser normal ou baixo em alguns casos e no modelo difícil o desempenho é maioritariamente baixo ou normal. Podemos observar que com o aumento da dificuldade os participantes acham que o seu desempenho decresce.

Grupo B No grupo B para cada um dos modelos e para cada uma das posições possíveis existem seis participantes.

Tabela 8.27: A exigência mental para este teste foi a adequada?

Modelo	P	Moda	Freq.
Fácil	1	CT	4
Fácil	2	CT	6
Fácil	3	CT	5
Médio	1	CT	3
Médio	2	CT	4
Médio	3	CT	5
Difícil	1	C	4
Difícil	2	CT	5
Difícil	3	CT	6

Tabela 8.28: A exigência física para este teste foi a adequada?

Modelo	P	Moda	Freq.
Fácil	1	CT	5
Fácil	2	CT	5
Fácil	3	CT	5
Médio	1	CT	5
Médio	2	CT/C	3/2
Médio	3	CT	6
Difícil	1	CT	5
Difícil	2	CT	5
Difícil	3	CT	4

Quase todos os participantes concordaram totalmente que a exigência mental do grupo B era a adequada para o teste realizado.

Mais uma vez neste grupo a exigência física consiste na capacidade de não se mexer bruscamente durante o teste, sendo que de forma constante todos concordaram que essa exigência era a adequada.

Tabela 8.29: A exigência temporal para este teste foi a adequada?

Modelo	P	Moda	Freq.
Fácil	1	CT	4
Fácil	2	CT	6
Fácil	3	CT	5
Médio	1	CT	5
Médio	2	CT	5
Médio	3	CT	5
Difícil	1	CT	5
Difícil	2	CT	6
Difícil	3	CT	5

Tabela 8.30: Como considera o seu esforço aplicado durante o teste?

Modelo	P	Moda	Freq.
Fácil	1	MB/B	2/2
Fácil	2	MB/B	2/2
Fácil	3	MB/N	3/2
Médio	1	MB/E	2/2
Médio	2	N/MB	3/2
Médio	3	MB/N/E	2/2/2
Difícil	1	B/N	3/3
Difícil	2	N	3
Difícil	3	N/B	3/2

A exigência temporal também é unanimemente favorável a todos os participantes. Este grupo em média foi o mais rápido e teoricamente mais fácil de concluir.

Em relação ao esforço aplicado durante o teste, os participantes de um modo geral concordaram que foi baixo ou muito baixo nos níveis fácil e médio, enquanto que no nível difícil o esforço já foi normal.

Tabela 8.31: Qual o seu nível de frustração durante o teste?

Modelo	P	Moda	Freq.
Fácil	1	MB/N	2/2
Fácil	2	MB	6
Fácil	3	MB/B	3/2
Médio	1	MB/B	2/2
Médio	2	B/MB	3/2
Médio	3	MB	4
Difícil	1	MB	4
Difícil	2	MB/B	2/2
Difícil	3	B/N	3/2

Tabela 8.32: Como classificaria a sua performance no teste?

Modelo	P	Moda	Freq.
Fácil	1	N	3
Fácil	2	N	5
Fácil	3	N	4
Médio	1	B/N	2/2
Médio	2	N	5
Médio	3	B/E	2/2
Difícil	1	N	4
Difícil	2	N	4
Difícil	3	N	4

O nível de frustração ao longo do teste e independentemente da ordem dos modelos, manteve-se constantemente baixa ou muito baixa.

Quanto à classificação do desempenho os participantes de um modo geral classificaram o seu desempenho como normal, tirando a exceção do nível médio em que o desempenho foi considerado como baixo ou elevado.

Grupo C No grupo C para cada um dos modelos e para cada uma das posições possíveis existem seis participantes.

Tabela 8.33: A exigência mental para este teste foi a adequada?

Modelo	P	Moda	Freq.
Fácil	1	CT	5
Fácil	2	CT	5
Fácil	3	CT	6
Médio	1	CT	5
Médio	2	CT/C	3/3
Médio	3	CT/C	3/2
Difícil	1	C	4
Difícil	2	CT	5
Difícil	3	CT/C	3/3

Tabela 8.34: A exigência física para este teste foi a adequada?

Modelo	P	Moda	Freq.
Fácil	1	CT	6
Fácil	2	CT	6
Fácil	3	CT	5
Médio	1	CT	6
Médio	2	CT	5
Médio	3	CT	5
Difícil	1	CT	5
Difícil	2	CT	6
Difícil	3	CT	6

Quanto à exigência mental exigida ao participante, estes concordaram que foi adequada em diferentes graus, ou seja, com o aumento dos níveis de dificuldade os participantes deixaram de concordar totalmente para passar a concordar em alguns casos.

Mais uma vez todos os participantes de forma uniforme concordam totalmente com a exigência física dos testes.

Tabela 8.35: A exigência temporal para este teste foi a adequada?

Modelo	P	Moda	Freq.
Fácil	1	CT	6
Fácil	2	CT	5
Fácil	3	CT	5
Médio	1	CT	5
Médio	2	CT	6
Médio	3	CT	4
Difícil	1	CT	5
Difícil	2	CT	5
Difícil	3	CT	6

Tabela 8.36: Como considera o seu esforço aplicado durante o teste?

Modelo	P	Moda	Freq.
Fácil	1	MB	3
Fácil	2	N/MB	3/2
Fácil	3	N	5
Médio	1	N/E	3/2
Médio	2	N	2
Médio	3	N/E	3/2
Difícil	1	N/E	3/3
Difícil	2	N	4
Difícil	3	E	4

De um modo muito uniforme os participantes acharam que a exigência temporal de todos os níveis de dificuldade dos modelos é adequada para o teste e a sua finalidade.

Neste grupo, ao contrário dos restantes existe uma alteração mais expressiva na alteração do esforço por parte dos participantes. À medida que o nível de dificuldade aumenta o esforço que o participante coloca no modelo em causa aumenta.

Tabela 8.37: Qual o seu nível de frustração durante o teste?

Modelo	P	Moda	Freq.
Fácil	1	MB	3
Fácil	2	MB	4
Fácil	3	MB/N	3/2
Médio	1	MB/N	2/2
Médio	2	B	3
Médio	3	E/N	3/2
Difícil	1	E/N	3/2
Difícil	2	N	3
Difícil	3	MB	3

Tabela 8.38: Como classificaria a sua performance no teste?

Modelo	P	Moda	Freq.
Fácil	1	N/B	3/2
Fácil	2	N	4
Fácil	3	N	4
Médio	1	B/MB	3/2
Médio	2	N/B	3/2
Médio	3	B/N	3/2
Difícil	1	N	5
Difícil	2	MB/B	3/3
Difícil	3	N/B	3/2

Em média o nível de frustração dos participantes é normal, muito baixo ou em casos esporádicos é elevado, como quando o modelo difícil é o primeiro a ser analisado.

A opinião do desempenho por parte dos participantes é maioritariamente normal quando o modelo é fácil, para baixo quando é o modelo médio e quando o modelo se torna difícil oscila entre muito baixo e normal. Existe uma pequena diminuição da confiança do trabalho feito por parte do participante com o aumento de dificuldade dos modelos do teste.

8.3 Hipóteses

Para poder responder às hipóteses anteriormente apresentadas é necessário comparar e analisar os dados apresentados nas tabelas deste capítulo e ainda analisar os elementos existentes na modelação dos modelos.

Devido à grande quantidade de hipóteses definidas e à escassez de tempo, foram selecionadas algumas hipóteses para serem detalhadas e apresentadas, sendo que as restantes serão analisadas e apresentadas em trabalhos de investigação posteriores.

8.3.1 $H_{confianca}$

Olhando em particular para hipótese da confiança, esta relaciona-se com o sentimento de uma maior taxa de sucesso ao longo do decorrer do teste. Esta pode ser analisada através da obtenção de dados do teste NASA-TLX com maiores graus de satisfação quando o teste avança e ao mesmo tempo sem que exista uma diminuição da percentagem de respostas corretas.

Tabela 8.39: Testes com aumento da taxa de sucesso e os seus respetivos dados performance NASA-TLX

Modelo	Nº	1º	TLX6	2º	TLX6	3º	TLX6
O1	1	3	B	3	B	4	B
O2	1	0	N	3	N	3	E
O3	2	1	B	2	B	3,5	B
O4	9	1,4	B	2,3	B	4,1	N
O5	3	1,3	N	2	N	2,7	N
O6	5	1,6	B	2,8	B	4,8	B
Total	21	1,4	B	2,5	B	4,1	B

Por outro lado, quando o participante começa a responder às perguntas dos modelos demasiado depressa, a sua taxa de sucesso decresce, levando a ponderar que o utilizador pode sentir-se frustrado e desmotivado para analisar e compreender os modelos. Estes

dados aliam-se ainda às respostas dadas no NASA-TLX serem dadas sem grande importância, podendo este facto ser analisado com uma grande rapidez das respostas durante estes testes e as suas comparações entre os três modelos diferentes.

Menos de metade dos participantes dos testes conseguiu aumentar o seu sucesso ao longo dos testes. Este fenómeno acontece maioritariamente quando o nível dos modelos decresce ao longo do teste. Em média, a taxa de sucesso quando é favorável tende a aumentar linearmente. Destes participantes, oito já possuíam algum conhecimento básico dos modelos BPMN devido a atividades académicas. Quarenta e quatro por cento das pessoas com conhecimento académico conseguiram ter um aumento da taxa de sucesso progressivo ao longo dos modelos, podendo este ser derivado da recordação progressiva dos conceitos. Mesmo que existam pessoas com algum conhecimento, no global todos os participantes com melhorias de performance não sentem confiança nas suas respostas. Em geral a mediana ou moda para os resultados em número par, a performance é classificada como baixa.

Com estes dados podemos concluir, que apesar de cerca de metade dos participantes conseguirem melhorar a sua performance, esta não se reflete na sua confiança na realização dos testes. Podemos então afirmar que a H1 desta hipótese não é verificada, mas este fenómeno não confirma a H0, para tal é preciso perceber qual é a frustração dos participantes que não melhoram as suas taxas de sucesso.

Tabela 8.40: Testes com diminuição da taxa de sucesso e os seus respetivos dados de frustração NASA-TLX

Modelo	Nº	1º	TLX5	2º	TLX5	3º	TLX5
O1	8	4,1	B	2,4	B	1,9	E
O2	8	3,8	MB	2,3	B	2,5	B
O3	7	2	MB	4,1	MB	2,3	B
O4	1	4	N	1	E	4	B
O5	6	2,2	N	3,2	MB	1,5	B
O6	4	2,3	N	1,8	B	3,5	B
Total	34	3,1	N	2,4	B	2,6	B

Ao nível da frustração dos participantes, é raramente elevada e por norma encontra-se sempre baixa ou normal, em contra partida os participantes de modo geral acham sempre que a sua performance é baixa ou normal. Quando os modelos mais fáceis são os primeiros a ser analisados a frustração do participante cresce e o contrário acontece quando os modelos mais fáceis são os últimos a analisar.

Em relação à H0 esta não é verificada no seu extremo, mas existe sim um aumento de frustração com o aumento do nível de dificuldade dos modelos e ao mesmo tempo existe um decréscimo da taxa de respostas corretas por modelo.

8.3.2 $H_{prestacao}$

Em relação à hipótese da perspectiva de prestação durante os testes por parte dos participantes, podemos afirmar que na sua grande maioria os participantes acham que tiveram uma performance normal ou baixa na análise dos modelos. Por norma, os participantes aproximam-se da realidade uma vez que existem poucos modelos com uma taxa de sucesso de 100% nas suas perguntas, mas por outro lado os participantes com melhores resultados não aumentam o seu nível de performance pois não têm confiança nas suas respostas, devido ao facto de poucos conhecerem os modelos e quando tal acontece o seu conhecimento não é significativo na realização dos testes.

Por isso podemos dizer que a hipótese H1 se verifica, ou seja, o utilizador consegue ter uma noção da sua prestação nos modelos de um modo geral.

8.3.3 $H_{PParticipantes}$

Para se poder perceber a compreensão dos papéis dos participantes dos modelos por parte dos seus utilizadores, deve-se analisar a performance dos participantes dos testes em relação às perguntas relacionadas com os participantes assim como relacionar as suas respostas com os seus tempos de resposta.

Tabela 8.41: Perguntas relacionadas com os participantes de um modelo

Grupo	Pergunta	Modelo	%C	Duração
A	Q1	Fácil	84%	00:38min
	Q1	Médio	57%	00:53min
	Q1	Difícil	68%	01:13min
A	Q2	Fácil	85%	00:25min
	Q2	Médio	58%	00:34min
	Q2	Difícil	57%	00:53min
A	Q4	Fácil	44%	00:50min
	Q4	Médio	6%	00:44min
	Q4	Difícil	43%	00:54min

Olhando para a tabela acima apresentada, podemos ver que por norma os utilizadores conseguem identificar os participantes dos processos modelados e relativamente rápido. Quanto à perceção dos participantes que obrigatoriamente participam nos processos é perceptível para metade dos participantes, ou seja, é confuso para metade dos utilizadores compreender a divisão de fluxos que pode ativar ou não a intervenção de participantes nos processos. Já na terceira pergunta, poucos participantes conseguiram perceber que os dados das atividades de um participante são os dados por ele utilizados.

Devido a estes resultados, podemos dizer que a H1 é verificada uma vez que os participantes dos testes perceberam quais os participantes dos processos e quando estes são ativos nos processos. Mas por outro lado, quando se fala na utilização de dados por parte das atividades a ser executadas pelos participantes a H0 é a que se aproxima mais da realizada.

8.3.4 $H_{\text{EfeitoAprendizagem}}$

Tal como apresentado nas primeiras hipóteses desta secção, é visível que existe um efeito de aprendizagem ao longo dos modelos. Vinte e um participantes aumentam o seu rendimento ao longo dos modelos, principalmente quando estes começam pelos modelos mais difíceis. Nos restantes participantes existem também taxas de crescimento entre os diferentes modelos, mas este não acontece entre todos os três modelos.

Podemos afirmar então, que a hipótese H1 se verifica nos modelos BPMN, isto deve-se ao facto de os utilizadores dos modelos começarem a perceber a lógica dos modelos e quais os alvos das perguntas colocadas aos modelos, uma vez que estas se mantêm inalteradas ao longo dos três modelos de um grupo.

8.4 Discussão dos Resultados

De um modo geral, pode-se realizar poucos testes de efeito significativo relativamente à duração e correção dos testes, devido aos dados não seguirem uma distribuição normal. Apesar de tudo, não podemos afirmar de um modo geral que a posição dos modelos no teste não influencia a correção e a duração dos testes no global. Por outro lado, já existem algumas perguntas onde a correção/duração das respostas às perguntas depende da dificuldade do modelo em causa.

Das hipóteses avaliadas podemos afirmar que a confiança dos participantes nas suas respostas às perguntas não melhora ao longo do teste $H_{\text{confianca}}$. Isto pode-se dever ao facto das dificuldades dos modelos variar, apresentado modelações distintas de análise podendo levar o utilizador a confundir os objetivos das perguntas e a duvidar das suas respostas.

De um modo geral, os participantes possuem uma noção realista da sua prestação nos testes $H_{\text{prestacao}}$. Esta perceção pode ser real, ou pode-se dever à falta de confiança nas suas respostas, em seguimento da hipótese anterior.

No que toca aos participantes de um modelo $H_{\text{PParticipantes}}$, o utilizador possui de um modo geral uma capacidade de perceção do que representa um participante no modelo. Por outro lado, os participantes não perceberam de um modo geral a existência de um participante “sistema” é um participante. Este facto pode estar relacionado com conhecimento prévio de outros diagramas de modelação que excluem este género de participantes e não é contabilizado como participante de um processo.

Relativamente ao efeito de aprendizagem dos modelos $H_{EfeitoAprendizagem}$, verificou-se mas não foi possível provar através das análises estatísticas. Este efeito pode ser derivado dos testes, mas foi combativo com a existência de várias ordens de aparecimento dos modelos de diferentes níveis de dificuldade. Por outro lado, o efeito de aprendizagem pode significar que os modelos são perceptíveis e mais fáceis de ler e analisar à medida que o contacto com os modelos aumenta.

De um modo geral, como estes modelos são usados por um grande leque de profissionais e áreas de trabalho, a existência de um efeito de aprendizagem e uma noção realista do que conseguiram perceber é uma vantagem e mais valia para estes modelos. Por outro lado, quando se trata de conhecimento específico da linguagem de modelação como os participantes é preciso ter cuidado e proceder a um estudo prévio.

8.4.1 Ameaças à validade

Em relação à validade das conclusões é preciso ter em conta que apesar de se ter recolhido uma quantidade bastante aceitável de participantes para os testes e estes participantes serem de diferentes anos de curso e conseqüentemente terem diferentes conhecimentos, existe sempre o risco da população testada não corresponder à população real. Para tal, foi iniciado o estudo com participantes profissionais sendo que estes devem também ter diferentes conhecimentos e percursos profissionais.

Na validação interna foi tido em conta o efeito de aprendizagem dos participantes após a conclusão de um teste, o que levou à restrição de participantes, ou seja, cada voluntário realizou um único teste disponível. Por outro lado, foi dada especial atenção ao ajuste correto dos modelos dos testes ao ecrã de teste e conseqüentemente aos tamanhos dos seus elementos e respetivas etiquetas.

Na validade externa conseguimos garantir que os participantes faziam parte do grupo alvo dos modelos **BPMN**, ou seja, pessoas com pouco conhecimento sobre as suas modelações e regras e por outro lado pessoas com conhecimento, que por norma são os seus criadores dos modelos. Este último grupo, ainda tem os seus testes a decorrer.

Por fim no que toca à validade construtiva, foi previamente explicado a todos os participantes que não estavam a ser avaliados, mas sim os modelos dos testes. Os participantes foram informados previamente do conceito do teste e os dos seus fluxos, ou seja, a existência de modelos, testes de usabilidade e um questionário demográfico. No entanto os participantes para além dos dados a serem recolhidos pelo *eye-tracker* não possuíam qualquer conhecimento sobre a lógica de recolha de dados do *framework*.

8.4.2 Inferências

Ao longo dos testes, pode-se dizer que era expectável que a posição e dificuldade dos modelos no teste influencia-se os resultados destes, assim como o tempo de resposta a cada pergunta dos diferentes modelos. Aliando a posição e a dificuldade dos modelos com o tempo de resposta às perguntas dos modelos e a correção das respostas dadas por parte

do utilizador, pode-se analisar e ganhar uma melhor perceção à cerca da facilidade de compreensão destes modelos por parte dos participantes, que por sua vez representa uma grande parte dos seus utilizadores.

No que toca à correção das respostas não é possível aferir num efeito significativo em relação à posição dos modelos ou da dificuldade dos modelos. Devendo-se ao facto dos dados não seguirem uma distribuição normal. Isto deve-se ao facto de existirem diferentes níveis de conhecimento dos participantes e estes possuírem modos de aprendizagem diferentes, assim como conhecimento da área de [ES](#) diferentes.

No que toca à duração de cada pergunta dos modelos dos testes, já existem algumas perguntas que possui uma distribuição normal, mas na grande maioria dos casos o efeito da posição ou dificuldade não é significativo. Quando os dados possuem um efeito significativo este não é relevante em termos quantitativos.

Apesar dos dados estatísticos não apresentarem muitos efeitos significativos, estes dão-nos a entender que estes modelos transmitem alguma confiança ao utilizador depois de alguma interação com estes, mas esta confiança nem sempre se reflete nos seus resultados. É perceptível ainda que muitos dos participantes se basearam em conhecimentos antigos de outros géneros de diagramas de modelação, podendo induzir o utilizador em erro por falta de conhecimento específico sobre os modelos [BPMN](#).

8.5 Sumário

Neste capítulo são apresentados os dados estatísticos tanto dos modelos como dos participantes e conseqüentemente, quais as inferências que podem ser feitas através deles. É ainda apresentada uma conclusão para as hipóteses apresentadas e no final é apresentado um resumo das conclusões retiradas da respetiva análise dos dados.

CONCLUSÕES

Neste capítulo é apresentada a síntese de todo o trabalho realizado nesta dissertação, assim como as suas principais contribuições e limitações sentidas. São apresentadas ainda as propostas de trabalho futuro e o trabalho que continua a decorrer no seguimento da dissertação.

9.1 Síntese do trabalho realizado

Com esta dissertação pretendeu-se avaliar a complexidade e a compreensibilidade dos modelos **BPMN**. Para tal foi desenvolvido um *framework* de avaliação, com duas grandes funcionalidades.

Para a avaliação dos modelos, foram propostas duas validações empíricas diferentes, a primeira consiste na análise estatística dos modelos e dos seus elementos, enquanto que a segunda consiste na análise em tempo real da interpretação e análise dos modelos por parte de voluntários.

As duas grandes funcionalidades do protótipo estão relacionadas com as duas validações empíricas dos modelos, ou seja, a primeira funcionalidade desenvolvida consiste na submissão de um modelo no *framework* e posteriormente o utilizador recebe os resultados no momento. A segunda funcionalidade consiste na realização dos testes, ou seja, apresentar os modelos e as respetivas questões ao utilizador de modo a recolher as suas respostas, por outro lado, é utilizado o *eye-tracker* a correr o seu *framework* em paralelo com o protótipo para recolher os dados dos seus dados faciais.

A análise dos modelos em si, foi realizada aos modelos dos testes realizados com os participantes para efeitos de comparação, mas pode ser realizado a qualquer modelo **BPMN** submetido na ferramenta. No que diz respeito aos testes com voluntários, foram reunidos cinquenta e cinco testes válidos com participantes com pouco ou nenhum conhecimento

sobre os modelos.

Após a recolha dos dados, pode ser iniciada a análise e comparação dos dados existentes. Sendo esta feita com recurso a várias técnicas de análise, como as análises descritivas, como por exemplo, médias e modas. Foram também analisados os dados dos utilizador com recurso à ferramenta SPSS. Esta permite-nos avaliar a normalidade dos dados, que por sua vez nos dá a informação necessária para proceder a análises do efeito significativo das variáveis testadas.

9.2 Contribuições

As principais contribuições desta dissertação consistem na implementação e recolha de métricas para avaliação dos modelos e a disponibilização da plataforma *online*, para ser acessível a qualquer pessoa interessada.

A ferramenta foi implementada com recurso a várias ferramentas e metodologias de desenvolvimento atuais e aceites pelos profissionais, tais como, o *Spring* como ligação à base de dados *SQL*, *React* como lógica de funcionamento *front-end*, *webservices REST* como *API's* de comunicação entre as diferentes camadas e *OCL* como definição de métricas.

Com a criação do protótipo e a realização dos testes, foi possível realizar a avaliação dos modelos e com isso contribuir com uma análise de uma amostra de participantes reais dos modelos. Esta permitiu perceber quais os elementos com pior interpretação nos modelos e a capacidade de aprendizagem dos mesmos.

9.3 Limitações

A grande limitação da dissertação foi a dificuldade da realização de testes com profissionais, devido às suas agendas e complicações de disponibilização do protótipo *online*.

Apesar de se ter conseguido obter uma amostra diversificada de participantes com poucos conhecimentos sobre os modelos, existe a limitação destes não conseguirem representar todos os perfis de utilizadores destes modelos.

Em relação ao *eye-tracker* em si, existe a limitação dos testes obrigatoriamente terem que ser realizados sempre num ecrã com o máximo de vinte e duas polegadas, o que reduz o tamanho máximo dos modelos que podem ser testados. Existe ainda a limitação dos erros introduzidos pelo dispositivo, que podem chegar a adicionar ou remover até meio centímetro da direção real do olhar dos participantes.

9.4 Trabalho Futuro

O trabalho futuro passa na continuação dos testes a decorrer com participantes com um conhecimento expressivo dos modelos, para que seja possível relacionar e perceber as diferenças entre os dois grupos de participantes e utilizadores dos modelos.

Um outro trabalho de futuro próximo é a continuação da análise de dados para que seja possível responder a novas hipóteses que surjam e sejam interessantes de analisar.

Um outro trabalho interessante será realizar testes com modelos ainda mais complexos e com mais especificidade das áreas de interesse dos modelos, para compreender e detalhar todos os cenários possíveis da análise e compreensão dos diferentes elementos dos modelos. Para tal, deve-se recorrer a ecrãs de teste maiores e conseqüentemente na aquisição de um dispositivo *eye-tracker* com maior capacidade de recolha de dados.

No que toca ao protótipo, foram planeadas duas funcionalidades extra, por implementar no protótipo com o objetivo de melhorar a experiência do utilizador na plataforma, sendo estas:

1. Visualização do histórico dos seus modelos submetidos.
2. Visualização dos dados gerais de todas as análises dos modelos submetidos na plataforma.

BIBLIOGRAFIA

- [Abr01] F. B. Abreu. “Using OCL to formalize object oriented metrics definitions”. Em: *INESC, Software Engineering Group, Technical Report ES007/2001* (2001).
- [Alb14] I. A. Alberto. “Desenvolvimento de plataforma de Business Process Management sobre SharePoint”. Tese de mestrado. Faculdade de Ciências - Universidade de Lisboa, 2014.
- [Alm14] A. C. Almeida. “Avaliação de Qualidade em Modelos de Requisitos i*”. Tese de mestrado. Faculdade de Ciências e Tecnologia - UNL, set. de 2014.
- [And+03] T. Andrews, F. Curbera, H. Dholakia, Y. Golland, J. Klein, F. Leymann, K. Liu, D. Roller, D. Smith, S. Thatte et al. *Business process execution language for web services*. 2003.
- [Ant+11] A. Antonini, A. M. Ferreira, S. Morasca e G. Pozzi. “Software measures for business processes”. Em: *Advances in Databases and Information Systems*. 2011, pp. 11–22.
- [AN05] J. Arlow e I. Neustadt. *UML 2 and the unified process: practical object-oriented analysis and design*. 2nd. Pearson Education, 2005, p. 312. ISBN: 3-321-32127-8.
- [AC06] C. Aschwanden e M. Crosby. “Code scanning patterns in program comprehension”. Em: *Proceedings of the 39th hawaii international conference on system sciences*. 2006.
- [Aya12] S. Ayad. “A quality based approach for the analysis and design of Business Process models”. Em: *Research Challenges in Information Science (RCIS), 2012 Sixth International Conference on*. IEEE. 2012, pp. 1–5.
- [Bar+15] H. M. S. Barona et al. “Checking and improving business process models in BPMN2”. Tese de doutoramento. Instituto Universitário de Lisboa (ISCTE), 2015.
- [Bas+94] V. Basili, G. Caldiera e H. Rombach. “The Goal Question Metric Approach”. Em: *Encyclopedia of software engineering 2.1994* (1994), pp. 528–532.
- [Bed12] R. Bednarik. “Expertise-dependent visual attention strategies develop over time during debugging with multiple code representations”. Em: *International Journal of Human-Computer Studies* 70.2 (2012), pp. 143–155.

- [BT06] R. Bednarik e M. Tukiainen. “An eye-tracking methodology for characterizing program comprehension processes”. Em: *Proceedings of the 2006 symposium on Eye tracking research & applications*. ACM. 2006, pp. 125–132.
- [BT08] R. Bednarik e M. Tukiainen. “Temporal eye-tracking data: evolution of debugging strategies with multiple representations”. Em: *Proceedings of the 2008 symposium on Eye tracking research & applications*. ACM. 2008, pp. 99–102.
- [Ben+07] R. Bendraou, B. Combemale, X. Crégut e M.-P. Gervais. “Definition of an Executable SPEM 2.0”. Em: *Software Engineering Conference, 2007. APSEC 2007. 14th Asia-Pacific*. IEEE. 2007, pp. 390–397.
- [Bin+13] D. Binkley, M. Davis, D. Lawrie, J. I. Maletic, C. Morrell e B. Sharif. “The impact of identifier style on effort and comprehension”. Em: *Empirical Software Engineering* 18.2 (2013), pp. 219–276.
- [BPMed] C. BPM. *BPMN 2.0 Implementation Reference*. Comunda Docs, Acedido em Novembro de 2017. URL: <https://docs.camunda.org/manual/7.7/reference/bpmn20/>.
- [Bra05] D. M. Brandon. *Project management for modern information systems*. IGI Global, 2005. ISBN: 1-59140-693-5.
- [Bri+96] L. C. Briand, S. Morasca e V. R. Basili. “Property-based software engineering measurement”. Em: *IEEE transactions on software Engineering* 22.1 (1996), pp. 68–86.
- [BK87] F Brooks e H. Kugler. *No silver bullet*. April, 1987.
- [Bus+14] T. Busjahn, R. Bednarik e C. Schulte. “What influences dwell time during source code reading?: analysis of element type and frequency as factors”. Em: *Proceedings of the Symposium on Eye Tracking Research and Applications*. ACM. 2014, pp. 335–338.
- [Car05a] J. Cardoso. “Control-flow complexity measurement of processes and Weyerker’s properties”. Em: *6th International Enformatika Conference*. Vol. 8. 2005, pp. 213–218.
- [Car05b] J. Cardoso. “How to measure the control-flow complexity of web processes and workflows”. Em: *Workflow handbook 2005* (2005), pp. 199–212.
- [Car06] J. Cardoso. “Approaches to Compute Workflow Complexity”. Em: *The Role of Business Processes in Service Oriented Architectures*. Ed. por F. Leymann, W. Reisig, S. R. Thatte e W. van der Aalst. Dagstuhl Seminar Proceedings 06291. Dagstuhl, Germany: Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany, 2006. URL: <http://drops.dagstuhl.de/opus/volltexte/2006/821>.

- [Car07] J. Cardoso. “Business process quality metrics: Log-based complexity of workflow patterns”. Em: *OTM Confederated International Conferences “On the Move to Meaningful Internet Systems”*. Springer. 2007, pp. 427–434.
- [Car+06] J. Cardoso, J. Mendling, G. Neumann e H. A. Reijers. “A discourse on complexity of process models”. Em: *International Conference on Business Process Management*. Springer. 2006, pp. 117–128.
- [Cened] I. K. Center. *Exemplo de um modelo BPEL*. (Acedido em Fevereiro de 2018). URL: https://www.ibm.com/support/knowledgecenter/en/SSBJDG_6.2.0/com.ibm.btools.modeler.advanced.help.doc/doc/concepts/widmodeling/bpelpick.html.
- [Cha+16] M. Challenger, F. Erata, M. Onat, H. Gezgen e G. Kardas. “A Model-Driven Engineering Technique for Developing Composite Content Applications”. Em: *OASICS-OpenAccess Series in Informatics*. Vol. 51. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik. 2016.
- [Che08] C.-Y. Cheng. *Complexity and usability models for business process analysis*. The Pennsylvania State University, 2008.
- [CP08] C. Y. Cheng e V. Prabhu. “Complexity model for business process analysis”. Em: *IIE Annual Conference. Proceedings*. Institute of Industrial e Systems Engineers (IISE). 2008, p. 1683.
- [Cor14] A. C. e Correia. “Quality of process modeling using BPMN: a model-driven approach”. Tese de doutoramento. Faculdade de Ciências e Tecnologias - UNL, 2014.
- [CS90] M. E. Crosby e J. Stelovsky. “How do we read algorithms? A case study”. Em: *Computer* 23.1 (1990), pp. 25–35.
- [Cro+02] M. E. Crosby, J. Scholtz e S. Wiedenbeck. “The roles beacons play in comprehension for novice and expert programmers”. Em: *14th Workshop of the Psychology of Programming Interest Group*. 2002, pp. 58–73.
- [DB07] R. Davis e E. Brabander. *ARIS Design Plataform: Getting started with BPM*. Springer Science & Business Media, 2007. ISBN: 978-1-84628-612-4.
- [DS+14] B. De Smet, L. Lempereur, Z. Sharafi, Y.-G. Guéhéneuc, G. Antoniol e N. Habra. “Taupe: Visualizing and analyzing eye-tracking data”. Em: *Science of Computer Programming* 79 (2014), pp. 260–278.
- [Deb+10] N Debnath, C. Salgado, M. Peralta, D. Riesco e G. Montejano. “Optimization of the Business Process metrics definition according to the BPDM standard and its formal definition in OCL”. Em: *Computer Systems and Applications (AICCSA), 2010 IEEE/ACS International Conference on*. IEEE. 2010, pp. 1–8.
- [DTH01] M. Dumas e A. H. Ter Hofstede. “UML activity diagrams as a workflow specification language”. Em: *UML*. Vol. 2185. Springer. 2001, pp. 76–90.

- [Edm95] B. Edmonds. “What is Complexity?-The philosophy of complexity per se with application to some examples in evolution”. Em: *The evolution of complexity*. Kluwer, Dordrecht, 1995.
- [Nin] *Exemplo de um modelo Nintex Workflow*. (Acedido em Janeiro de 2018). URL: <http://www.enjoysharepoint.com/photo/nintex-workflow-assign-task-demo.png>.
- [Spe] *Exemplo de um modelo SPEM*. (Acedido em Fevereiro de 2018). URL: https://www.sparxsystems.com.au/enterprise_architect_user_guide/13.0/model_domains/spem_elements_and_connectors.html.
- [Fer17] P. V. Fernandes. “A Framework for Configuring and Specifying Configurable Business Process Models”. Tese de mestrado. Faculdade de Ciências e Tecnologia - UNL, dez. de 2017.
- [Fered] L. Fernando. *Exemplo de um modelo BPT*. (Acedido em Janeiro de 2018). URL: <https://leonardofernandes.outsystemscloud.com/eventshelp/img/lesson0207.png?1798>.
- [FL11] K. Figl e R. Laue. “Cognitive complexity in business process modeling”. Em: *International Conference on Advanced Information Systems Engineering*. Springer. 2011, pp. 452–466.
- [Mat] *Framework Materialize*. (Acedido em Setembro de 2018). URL: <https://materializecss.com/>.
- [Fri+14] T. Fritz, A. Begel, S. C. Müller, S. Yigit-Elliott e M. Züger. “Using physiological measures to assess task difficulty in software development”. Em: *Proceedings of the 36th international conference on software engineering*. ACM. 2014, pp. 402–413.
- [GRed] D. Gagné e S. Ringuette. *BPMN Quick Guide*. 2nd. Object Management Group, (Acedido em Novembro de 2017). URL: <http://www.bpmn.org/>.
- [GL06a] V. Gruhn e R. Laue. “Adopting the cognitive complexity measure for business process models”. Em: *Cognitive Informatics, 2006. ICCI 2006. 5th IEEE International Conference on*. Vol. 1. IEEE. 2006, pp. 236–241.
- [GL06b] V. Gruhn e R. Laue. “Complexity metrics for business process models”. Em: *9th international conference on business information systems (BIS 2006)*. Vol. 85. 2006, pp. 1–12.
- [GL07] V. Gruhn e R. Laue. “Approaches for business process model complexity metrics”. Em: *Technologies for Business Information Systems (2007)*, pp. 13–24.
- [GV04] P. Grunwald e P. Vitányi. “Shannon information and Kolmogorov complexity”. Em: *arXiv preprint cs/0410002 (2004)*.

- [Hen16] H. G. Henriques. “Domain Specific Language Evaluation: OutSystems’ Business Process Technology”. Tese de mestrado. Faculdade de Ciências e Tecnologia - UNL, 2016.
- [HVM11] J. M. Hernández e P. Van Mieghem. “Classification of graph metrics”. Em: *Delft University of Technology, Tech. Rep* (2011).
- [IEE98] IEEE. “IEEE Standard for a Software Quality Metrics Methodology”. Em: *IEEE-SA Standards Board* (1998). DOI: [10.1109/IEEESTD.1998.243394](https://doi.org/10.1109/IEEESTD.1998.243394). URL: <http://ieeexplore.ieee.org/servlet/opac?punumber=6061>.
- [Inc17] B. P. Incubator. *BPI Web Registration without Moderator*. 2017 (Acedido em Novembro de 2017). URL: <https://www.businessprocessincubator.com/content/bpi-web-registration-without-moderator/>.
- [Jea+09] S. Jeanmart, Y.-G. Gueheneuc, H. Sahraoui e N. Habra. “Impact of the visitor pattern on program comprehension and maintenance”. Em: *Empirical Software Engineering and Measurement, 2009. ESEM 2009. 3rd International Symposium on*. IEEE. 2009, pp. 69–78.
- [Jun+04] H.-W. Jung, S.-G. Kim e C.-S. Chung. “Measuring software product quality: A survey of ISO/IEC 9126”. Em: *IEEE software* 21.5 (2004), pp. 88–92.
- [Kea+86] J. P. Kearney, R. L. Sedlmeyer, W. B. Thompson, M. A. Gray e M. A. Adler. “Software complexity measurement”. Em: *Communications of the ACM* 29.11 (1986), pp. 1044–1050.
- [Khl+09] W. Khelif, L. Makni, N. Zaaboub e H. Ben-Abdallah. “Quality metrics for business process modeling”. Em: *Proceedings of the 9th WSEAS international conference on Applied computer science*. World Scientific, Engineering Academy e Society (WSEAS). 2009, pp. 195–200.
- [KN12] K. Kluza e G. J. Nalepa. “Proposal of square metrics for measuring business process model complexity”. Em: *Computer Science and Information Systems (FedCSIS), 2012 Federated Conference on*. IEEE. 2012, pp. 919–922.
- [LA09] K. B. Lassen e W. M. van der Aalst. “Complexity metrics for Workflow nets”. Em: *Information and Software Technology* 51.3 (2009), pp. 610–626.
- [LK01] A. Latva-Koivisto. “Finding a complexity measure for business process models”. Em: (2001).
- [Htm] *Linguagem HTML*. (Acedido em Setembro de 2018). URL: <https://www.w3schools.com/html/>.
- [Jav] *Linguagem JavaScript*. (Acedido em Setembro de 2018). URL: <https://www.w3schools.com/js/>.
- [Mak+10] L. Makni, W. Khelif, N. Z. Haddar e H. Ben-Abdallah. “A Tool for Evaluating the Quality of Business Process Models.” Em: *ISSS/BPSC*. 2010, pp. 230–242.

- [Men] J. Mendling. *Metrics for Business Process Models*. Vol. 6. Springer, Berlin, Heidelberg. ISBN: 978-3-540-89223-6. DOI: https://doi.org/10.1007/978-3-540-89224-3_4.
- [Men+07] J. Mendling, H. Reijers e J. Cardoso. “What makes process models understandable?” Em: *Business process management* (2007), pp. 48–63.
- [Muk08] G. Muketha. “Complexity Metrics for Measuring the Understandability and Maintainability of Business Process Models using Goal-Question-Metric (GQM)”. Em: (2008).
- [Muk10] G. Muketha. “A survey of business processes complexity metrics”. Em: *Information Technology Journal* 9 (2010), pp. 1336–1344.
- [Mur89] T. Murata. “Petri nets: Properties, analysis and applications”. Em: *Proceedings of the IEEE* 77.4 (1989), pp. 541–580.
- [OM08] J. A. O’Brien e G. M. Marakas. *Introduction to information systems*. 15th. McGraw-Hill/Irwin New York City, USA, 2008. ISBN: 978-0-07-337677-6.
- [OMG11] OMG. *Business Process Model and Notation*. 2011, (Acedido em Outubro de 2017). URL: <http://www.omg.org/spec/BPMN/2.0>.
- [OMG14] OMG. “Object Constraint Language Version 2.4”. Em: *OMG Specification, Object Management Group* (2014).
- [Ço14] E. Çoskun. “A new complexity metric for business process models”. Tese de mestrado. atilim university, jul. de 2014.
- [PM13] R. Petrusel e J. Mendling. “Eye-tracking the factors of process model comprehension tasks”. Em: *International Conference on Advanced Information Systems Engineering*. Springer. 2013, pp. 224–239.
- [Pin+12] J. Pinggera, M. Furtner, M. Martini, P. Sachse, K. Reiter, S. Zugal e B. Weber. “Investigating the process of process modeling with eye movement analysis”. Em: *International Conference on Business Process Management*. Springer. 2012, pp. 438–450.
- [PC17] G. Polančič e B. Cegnar. “Complexity metrics for process models—A systematic literature review”. Em: *Computer Standards & Interfaces* 51 (2017), pp. 104–117.
- [Pol13] G. Polancicl. *Understanding BPMN Connections*. Orbus Software, 2013. URL: http://www3.cis.gsu.edu/dtruex/courses/CIS4120/Sessions/Session%20-%20-%20Level%201%20BPMN/Introduction_to_BPMN/Understanding-the-bpmn-connections-Orbus.pdf.
- [PG10] G. C. Porras e Y.-G. Guéhéneuc. “An empirical study on the efficiency of different design pattern representations in UML class diagrams”. Em: *Empirical Software Engineering* 15.5 (2010), pp. 493–522.

- [RG02] M. Richters e M. Gogolla. “OCL: Syntax, semantics, and tools”. Em: *Object Modeling with the OCL*. Springer, 2002, pp. 42–68.
- [Rod+14] P. Rodeghero, C. McMillan, P. W. McBurney, N. Bosch e S. D’Mello. “Improving automated source code summarization via an eye-tracking study of programmers”. Em: *Proceedings of the 36th International Conference on Software Engineering*. ACM. 2014, pp. 390–401.
- [Rol+06] E. Rolón, F. Ruiz, F. García e M. Piattini. “Applying software metrics to evaluate business process models”. Em: *CLEI-Electronic Journal* 9.1 (2006).
- [Rus+06] N. Russell, W. M. van der Aalst, A. H. Ter Hofstede e P. Wohed. “On the suitability of UML 2.0 activity diagrams for business process modelling”. Em: *Proceedings of the 3rd Asia-Pacific conference on Conceptual modelling*. Vol. 53. Australian Computer Society, Inc. 2006, pp. 95–104.
- [SG+10] L. Sánchez González, F. García Rubio, F. Ruiz González e M. Piattini Velthuis. “Measurement in business processes: a systematic review”. Em: *Business Process Management Journal* 16.1 (2010), pp. 114–134.
- [SG+12a] L. Sánchez-González, F. García, F. Ruiz e J. Mendling. “A study of the effectiveness of two threshold definition techniques”. Em: (2012).
- [SG+12b] L. Sánchez-González, F. García, F. Ruiz e J. Mendling. “Quality indicators for business process models from a gateway complexity perspective”. Em: *Information and Software Technology* 54.11 (2012), pp. 1159–1174.
- [Sha+11] R. Shapiro, S. A. White, N. Palmer, M. zur Muehlen, T. Allweyer e D. Gagné. *BPMN 2.0 Handbook Second Edition: Methods, Concepts, Case Studies and Standards in Business Process Management Notation*. Future Strategies Inc., 2011. ISBN: 978-0-9819870-7-1.
- [Sha+13] Z. Sharafi, A. Marchetto, A. Susi, G. Antoniol e Y.-G. Gueheneuc. “An empirical study on the efficiency of graphical vs. textual representations in requirements comprehension”. Em: *Program Comprehension (ICPC), 2013 IEEE 21st International Conference on*. IEEE. 2013, pp. 33–42.
- [Sha+15] Z. Sharafi, Z. Soh e Y.-G. Guéhéneuc. “A systematic literature review on the usage of eye-tracking in software engineering”. Em: *Information and Software Technology* 67 (2015), pp. 79–107.
- [SM10] B. Sharif e J. I. Maletic. “An eye tracking study on the effects of layout in understanding the role of design patterns”. Em: *Software Maintenance (ICSM), 2010 IEEE International Conference on*. IEEE. 2010, pp. 1–10.
- [Sha+12] B. Sharif, M. Falcone e J. I. Maletic. “An eye-tracking study on the role of scan time in finding source code defects”. Em: *Proceedings of the Symposium on Eye Tracking Research and Applications*. ACM. 2012, pp. 381–384.

- [Soh+12] Z. Soh, Z. Sharafi, B. Van den Plas, G. C. Porras, Y.-G. Guéhéneuc e G. Antoniol. “Professional status and expertise for UML class diagram comprehension: An empirical study”. Em: *Program Comprehension (ICPC), 2012 IEEE 20th International Conference on*. IEEE. 2012, pp. 163–172.
- [Som16] I. Sommerville. *Software Engineering*. 10th. USA: Addison-Wesley Publishing Company, 2016. ISBN: 1-298-09613-6.
- [Spr13] C. Sprague. “Information Systems Development.” Em: *Research Starters: Business (Online Edition)* (2013).
- [Ste+] R. Steenweg, M. Kuhrmann e D. M. Fernández. “Software Engineering Process Metamodels”. Em: ().
- [Tur+14] R. Turner, M. Falcone, B. Sharif e A. Lazar. “An eye-tracking study assessing the comprehension of C++ and Python source code”. Em: *Proceedings of the Symposium on Eye Tracking Research and Applications*. ACM. 2014, pp. 231–234.
- [Uwa+06] H. Uwano, M. Nakamura, A. Monden e K.-i. Matsumoto. “Analyzing individual performance of source code review using reviewers’ eye movement”. Em: *Proceedings of the 2006 symposium on Eye tracking research & applications*. ACM. 2006, pp. 133–140.
- [Van+08] I. Vanderfeesten, H. A. Reijers, J. Mendling, W. M. van der Aalst e J. Cardoso. “On a quest for good process models: the cross-connectivity metric”. Em: *International Conference on Advanced Information Systems Engineering*. Springer. 2008, pp. 480–494.
- [Wes10] M. Weske. *Business Process Management: Concepts, Languages, Architectures*. Springer Publishing Company, Incorporated, 2010. ISBN: 978-3-540-73521-2.
- [Whi08] S. White. *BPMN modeling and reference guide: understanding and using BPMN*. Florida, USA: Future Strategies Inc., 2008.
- [Yus+07] S. Yusuf, H. Kagdi e J. I. Maletic. “Assessing the comprehension of UML class diagrams via eye tracking”. Em: *Program Comprehension, 2007. ICPC’07. 15th IEEE International Conference on*. IEEE. 2007, pp. 113–122.
- [Zag+16] J. Zagermann, U. Pfeil e H. Reiterer. “Measuring Cognitive Load using Eye Tracking Technology in Visual Computing”. Em: *BELIV’16: Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization*. 2016, pp. 78–85.



EXEMPLOS DOS MODELOS BPM

A.1 Diagramas de atividades

Na [Figura A.1](#) está representado um mecanismo de *login* numa plataforma, em que para isso é necessário fornecer alguns dados. Os dados inseridos pelo utilizador serão validados, passando assim a ser possível avançar para a autenticação na plataforma. Caso o utilizador deseje é possível cancelar a ação através de uma notificação sua, representada no diagrama através de um *signal*.

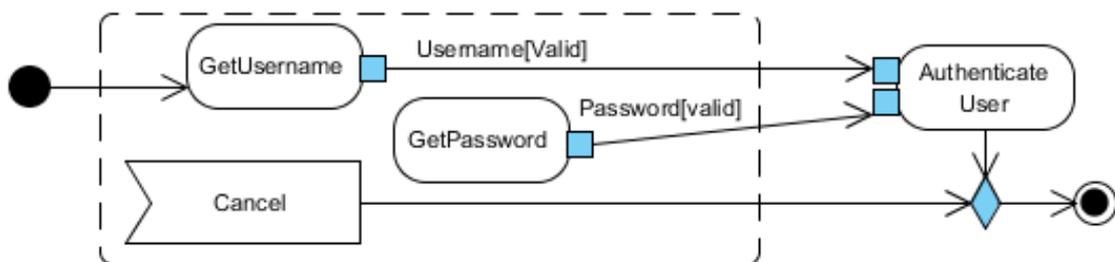


Figura A.1: Exemplo de um diagrama de atividades [AN05]

A.2 SPEM

Na Figura A.2 é possível verificar que existem dois participantes, sendo que ambos podem executar o processo em causa. O processo consiste numa sequência de atividades sobre o desenvolvimento de um *software*, e para tal gera dois documentos/dados distintos e necessita de dados gerados por outro processo.

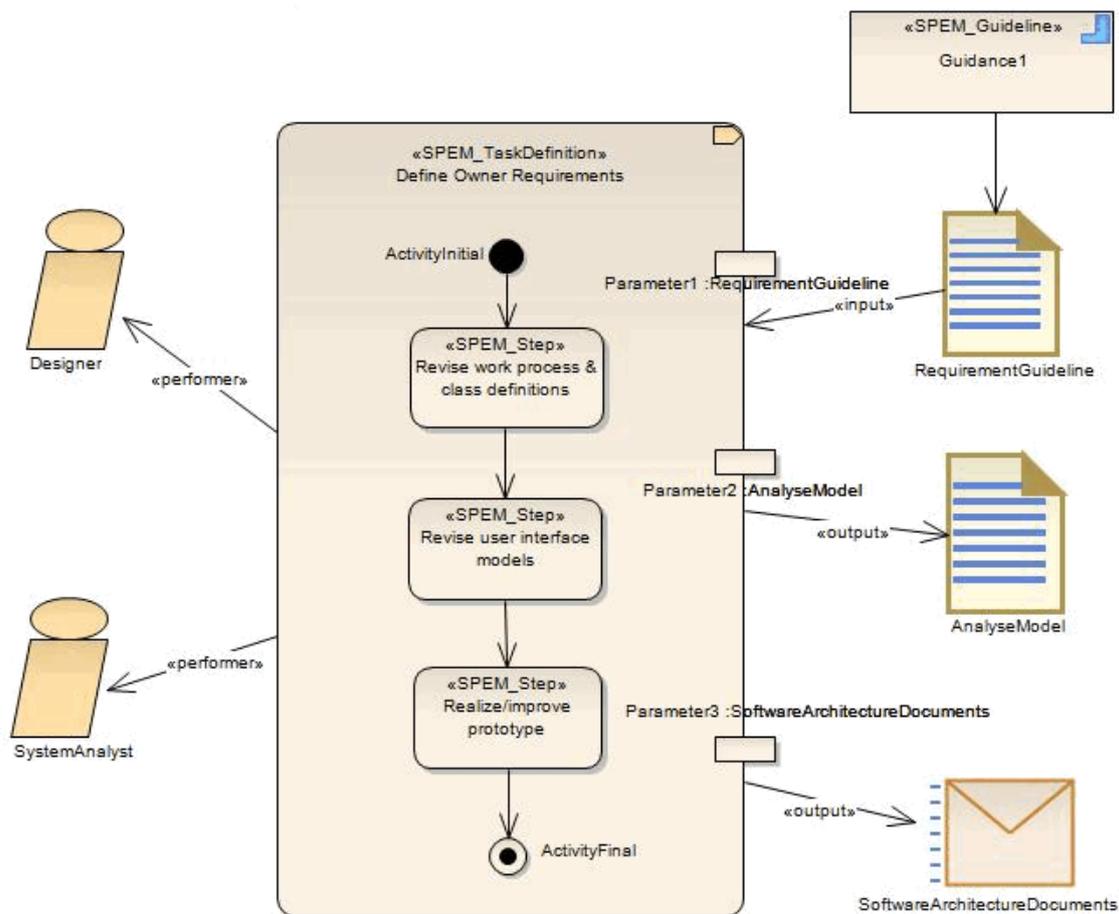


Figura A.2: Exemplo de um diagrama SPEM [Spe]

A.3 BPEL

Na Figura A.3 é representado o fluxo de um pagamento. Este pode ser feito de duas maneiras distintas, com cartão ou em dinheiro. Após a escolha o fluxo volta a unir-se e posteriormente a compra é dada como finalizada.

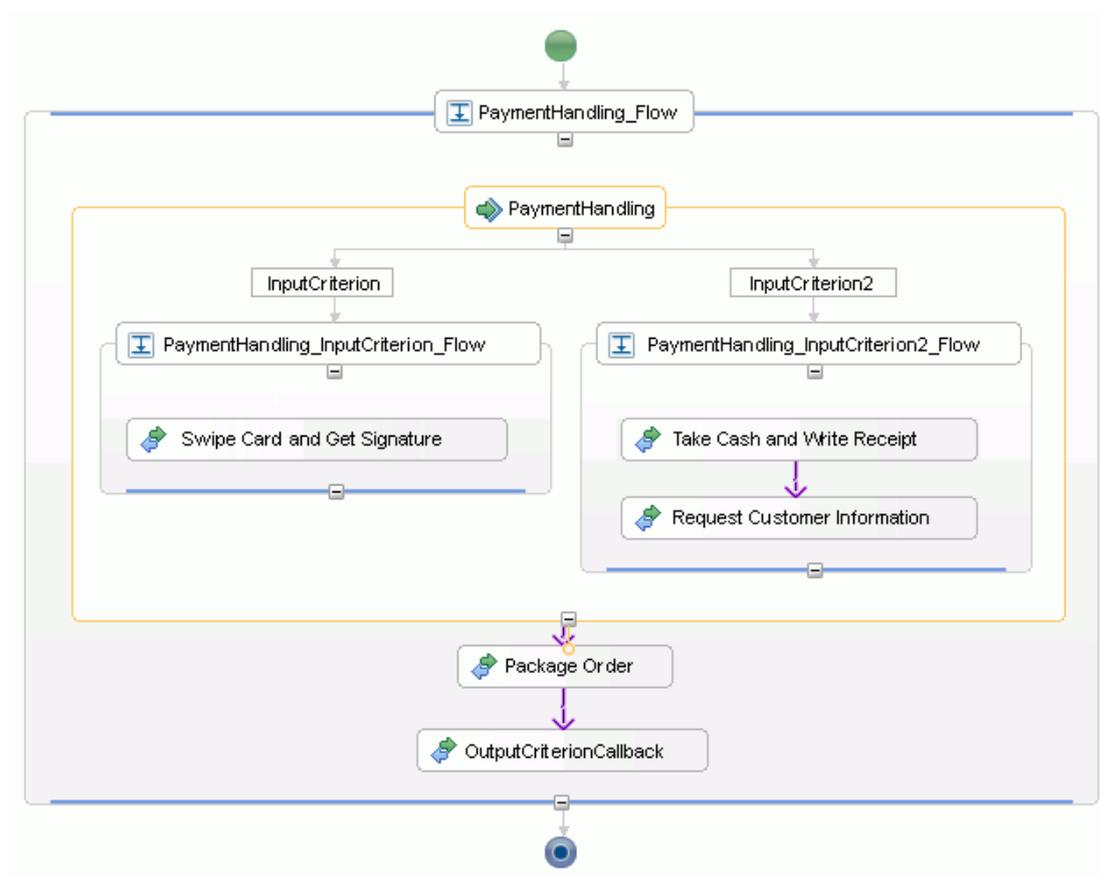


Figura A.3: Exemplo de um diagrama BPEL [Cened]

A.4 BPT

Na [Figura A.4](#) é apresentado um diagrama **BPT** que representa uma lógica de validação de um formulário. De um modo mais detalhado, quando o formulário está a ser validado, se este for válido é feita uma atualização dos campos modificados e o utilizador recebe uma notificação de modo a finalizar a ação. Caso o formulário não esteja correto o utilizador é notificado de quais os campos errados e o processo é terminado, podendo ser outra vez acionado por parte do utilizador.

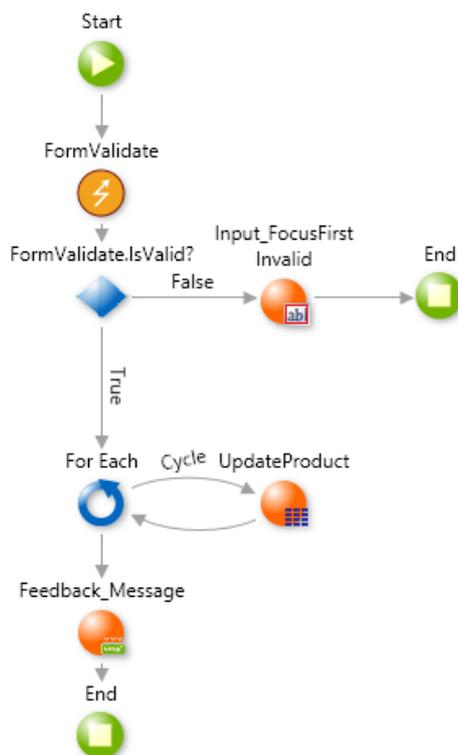


Figura A.4: Exemplo de um diagrama BPT [Fered]

A.5 Nintex Workflow

Na [Figura A.5](#) encontra-se um exemplo deste tipo de diagramas, cuja funcionalidade é resolver uma tarefa. A tarefa pode ser: aprovada, rejeitada ou cancelada. Cada uma destas decisões leva a uma entrada no histórico diferente.

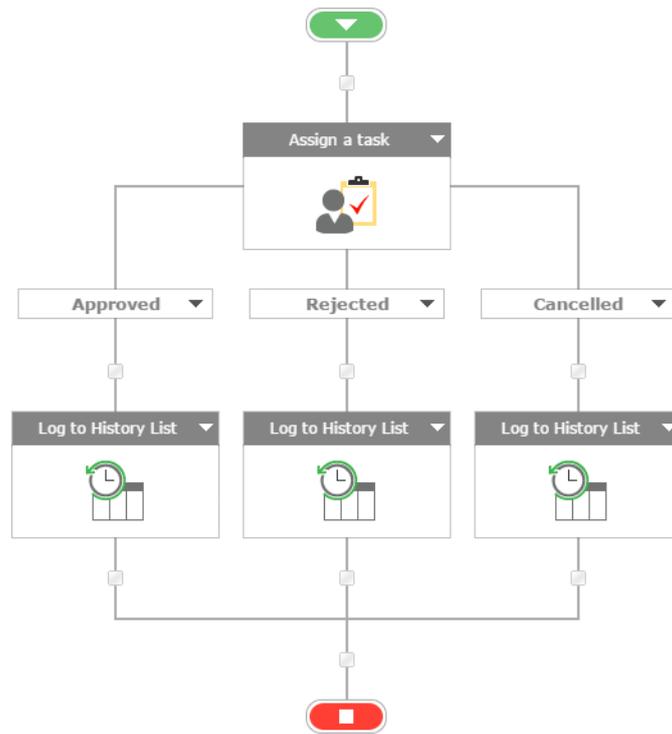


Figura A.5: Exemplo de um diagrama *Nintex workflow* [Nin]

EXEMPLO OCL

Na [Figura B.1](#) é apresentado um diagrama de classes, representado de uma maneira simplista os departamentos e empregados de uma empresa.

O Diagrama contém três diferentes classes: **Employee**, **Department**, **Project**. Cada uma destas classes contém os seus atributos com os seus tipos correspondentes. A classe **Employee** ainda apresenta uma função, com um parâmetro de entrada e tipo de retorno da função.

Em termos de associações, um empregado trabalha em um ou mais departamentos e um departamento pode conter diversos empregados, ou até nenhum. Um departamento contém diversos projetos. Um projeto por sua vez, têm vários empregados e um empregado pode estar associado a vários projetos.

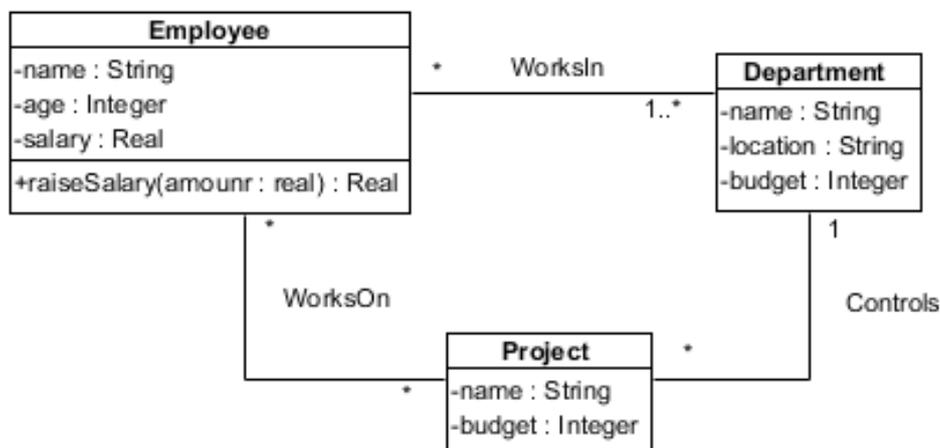


Figura B.1: Exemplo de um diagrama de classes [RG02]

Com este diagrama de classes, pode-se especificar em OCL as suas relações e restrições, apresentadas na [Listagem B.1](#) e [B.2](#). Primeiro são apresentadas as definições do diagrama de classes [Listagem B.1](#) e por último são apresentadas duas restrições e uma operação [Listagem B.2](#).

Listagem B.1: Definição do modelo em OCL [[RG02](#)]

```

1  -- model
2
3  model Company
4
5  class Employee
6  attributes
7    name : String
8    age : Integer
9    salary : Real
10 operations
11   raiseSalary(amount : Real) : Real
12 end
13
14 class Department
15 attributes
16   name : String
17   location : String
18   budget : Integer
19 end
20
21 class Project
22 attributes
23   name : String
24   budget : Integer
25 end
26
27 association WorksIn between
28   Employee[*]
29   Department[1..*]
30 end
31
32 association WorksOn between
33   Employee[*]
34   Project[*]
35 end
36
37 association Controls between
38   Department[1]
39   Project[*]
40 end

```

São definidas duas restrições e uma operação [Listagem B.2](#). A primeira restrição limita o valor mínimo de um salário, enquanto que a segunda nos dá a restrição sobre o montante do salário de um empregado, ou seja, um empregado integrado em mais projetos possui um ordenado mais elevado. A função possui uma pré-condição, que limita o aumento de ordenado a um valor positivo e a pós-condição diz-nos que o salário atual será o valor atual mais o valor do aumento.

Listagem B.2: Definição de restrições em OCL [[RG02](#)]

```
1 -- OCL constraints
2
3 constraints
4
5 context Department inv:
6   self.budget >= 0
7   self.employee->forall(e1, e2 | e1.project->size > e2.project->size)
8     implies e1.salary > e2.salary
9
10 context Employee::raiseSalary(amount : Real) : Real
11   pre: amount > 0
12   post: self.salary = self.salary@pre + amount and result = self.salary
```




MODELOS AVALIADOS

Os seguintes modelos representam os modelos de teste da validação empírica com participantes. Os modelos estão divididos por grupos e cada grupo contém três modelos de nível diferentes, apresentados de modo crescente quanto ao seu nível de dificuldade.

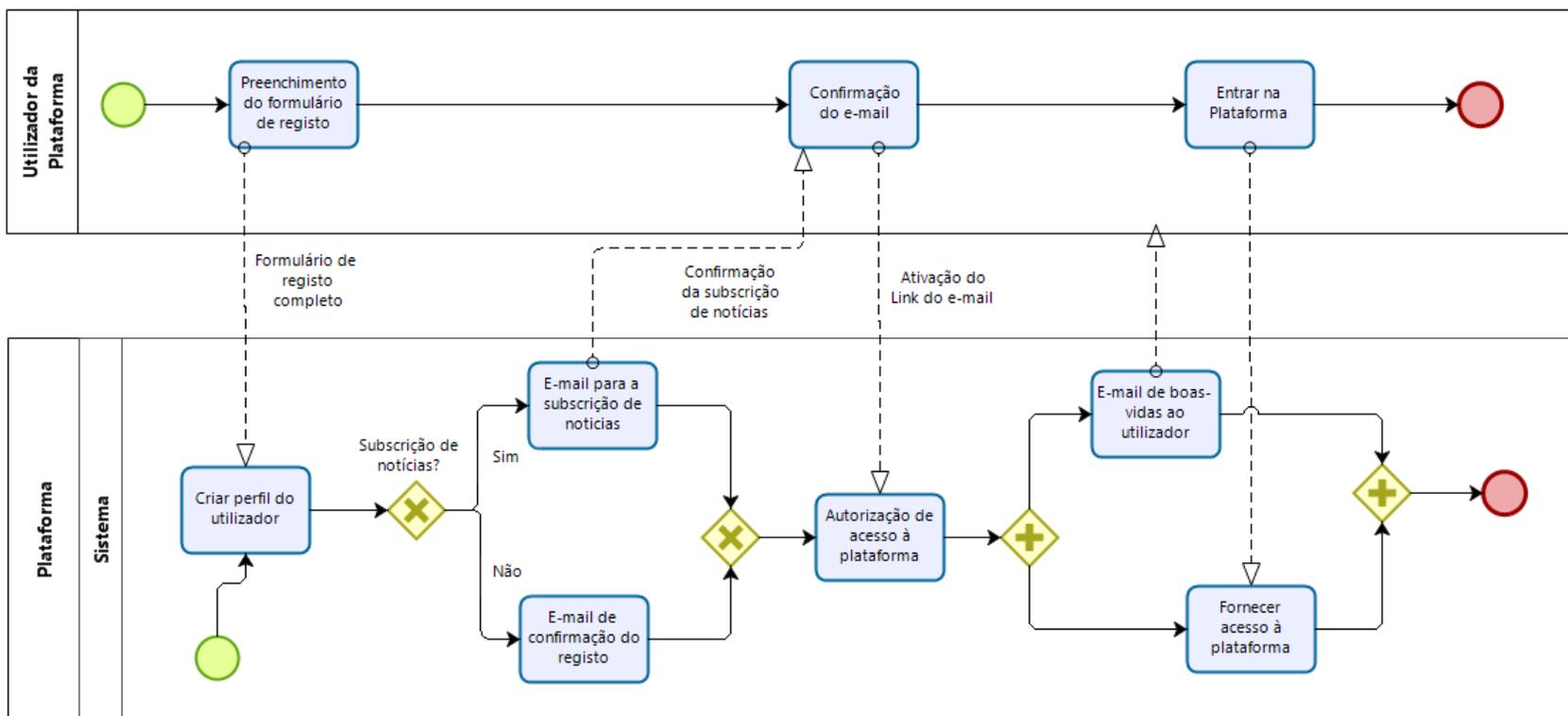


Figura C.1: Modelo fácil do grupo A

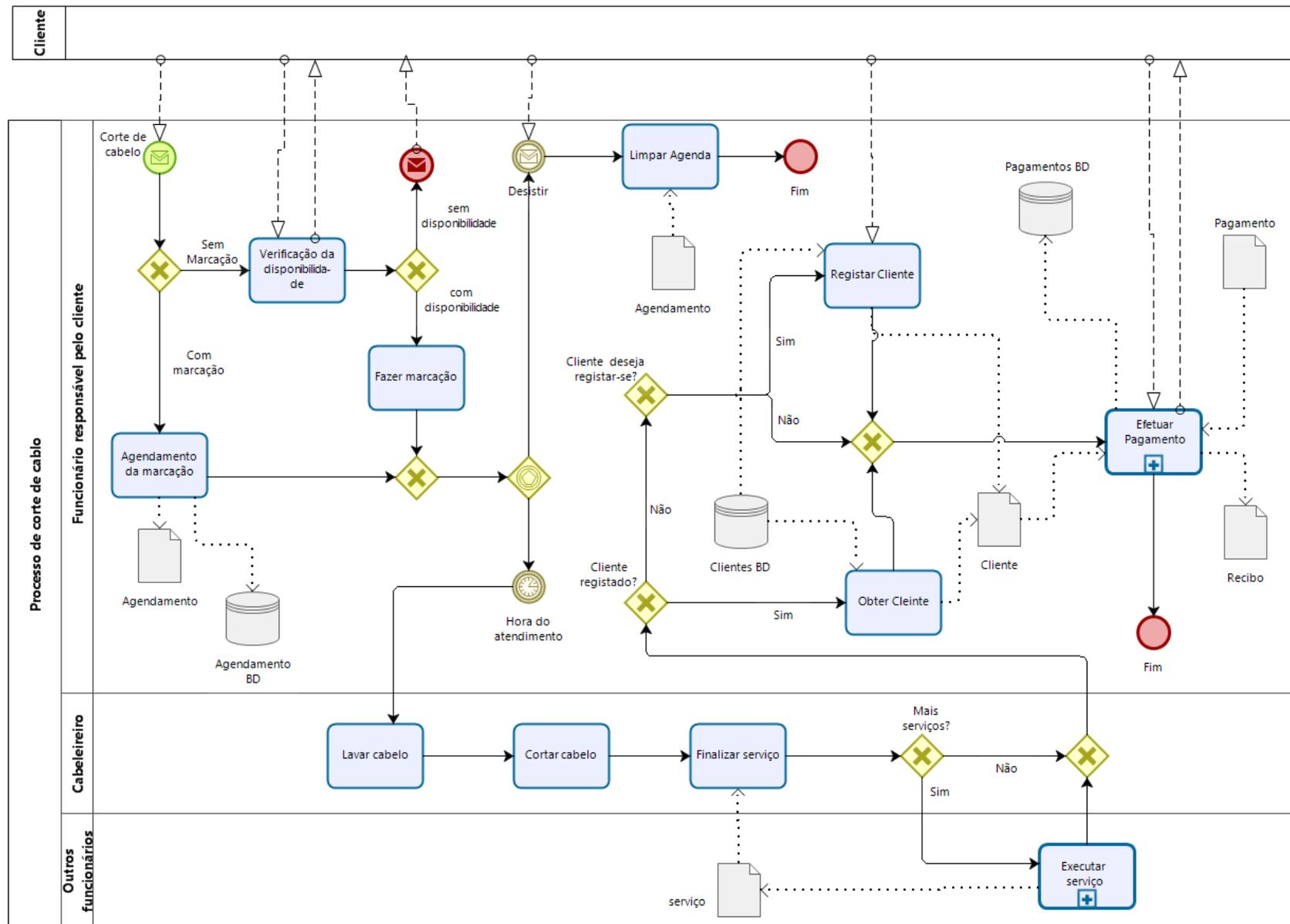


Figura C.2: Modelo intermédio do grupo A

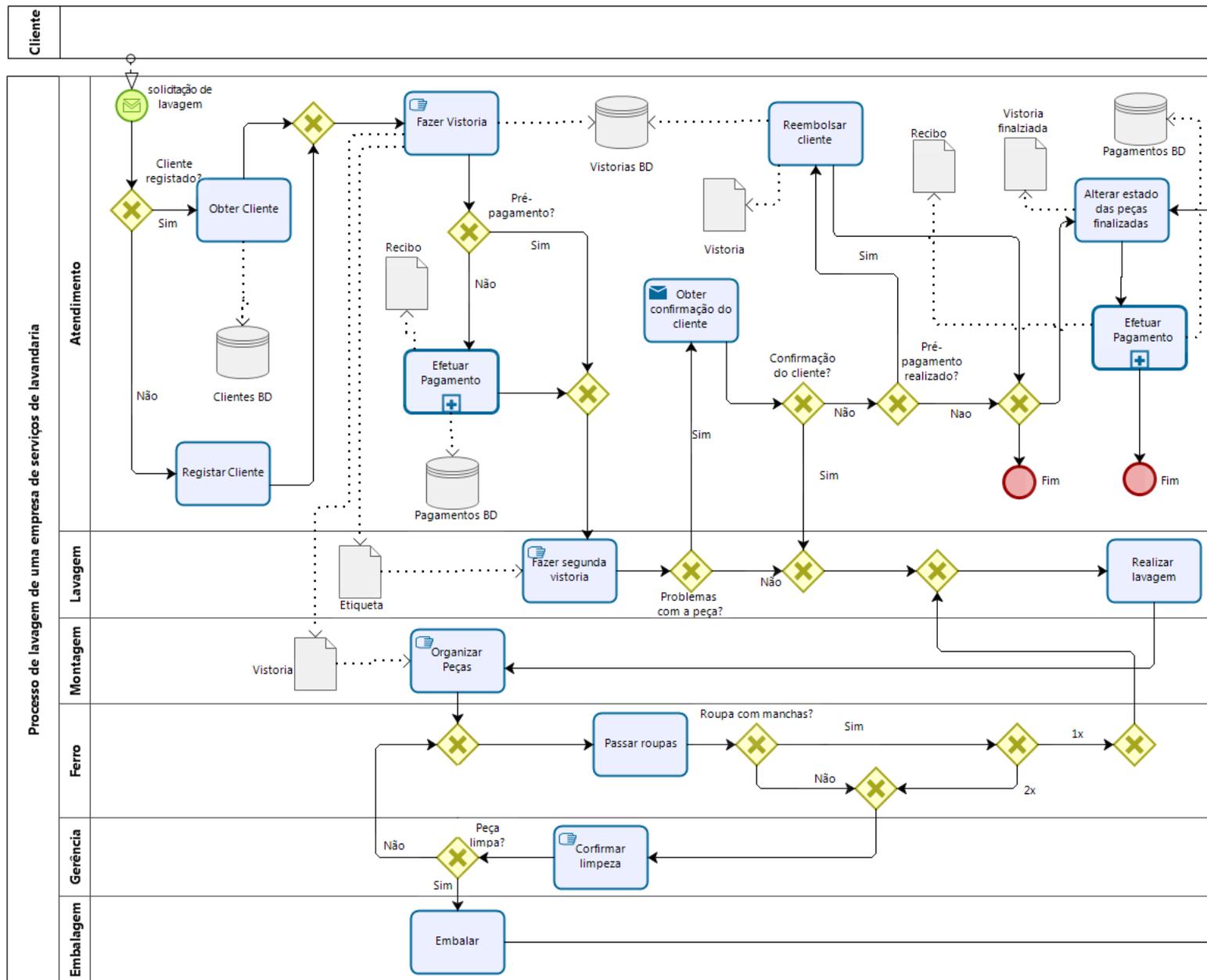


Figura C.3: Modelo difícil do grupo A

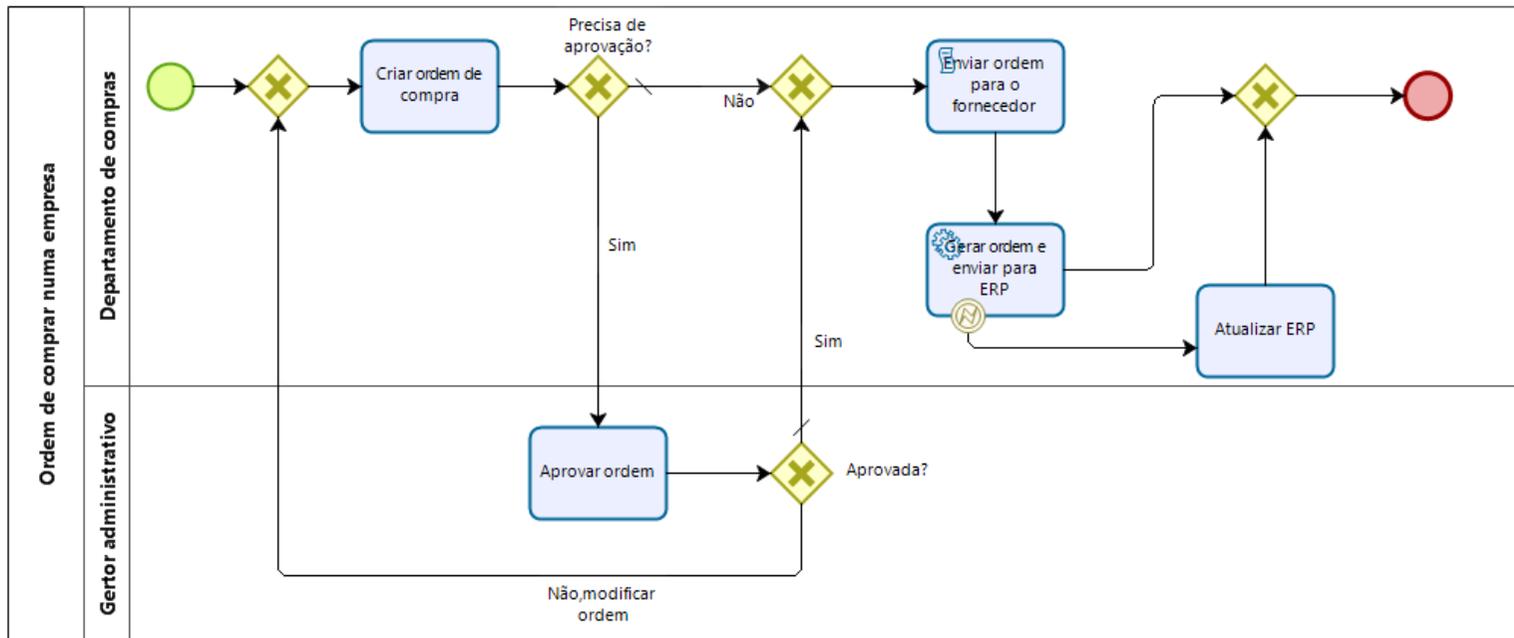


Figura C.4: Modelo fácil do grupo B

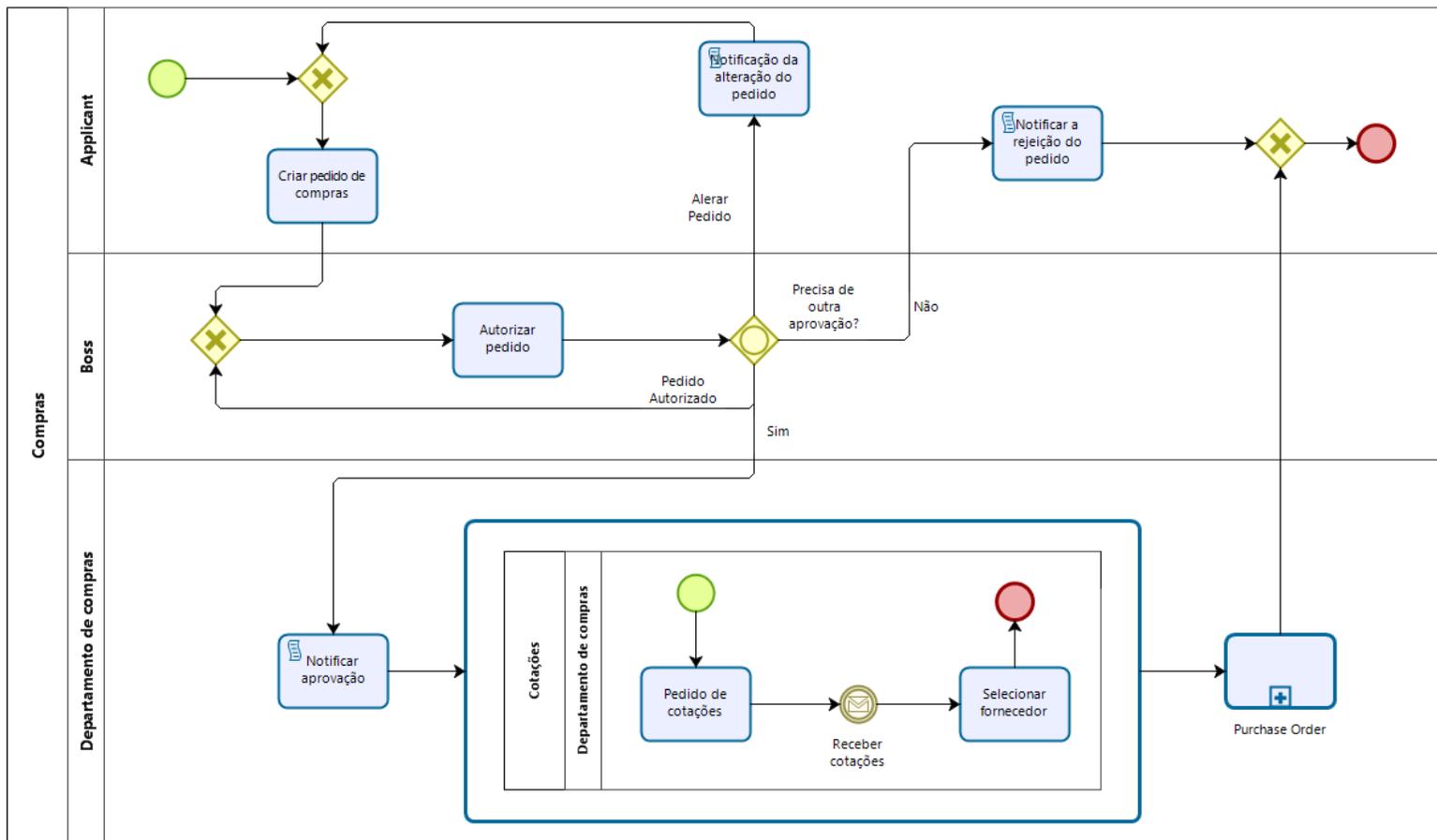


Figura C.5: Modelo intermédio do grupo B

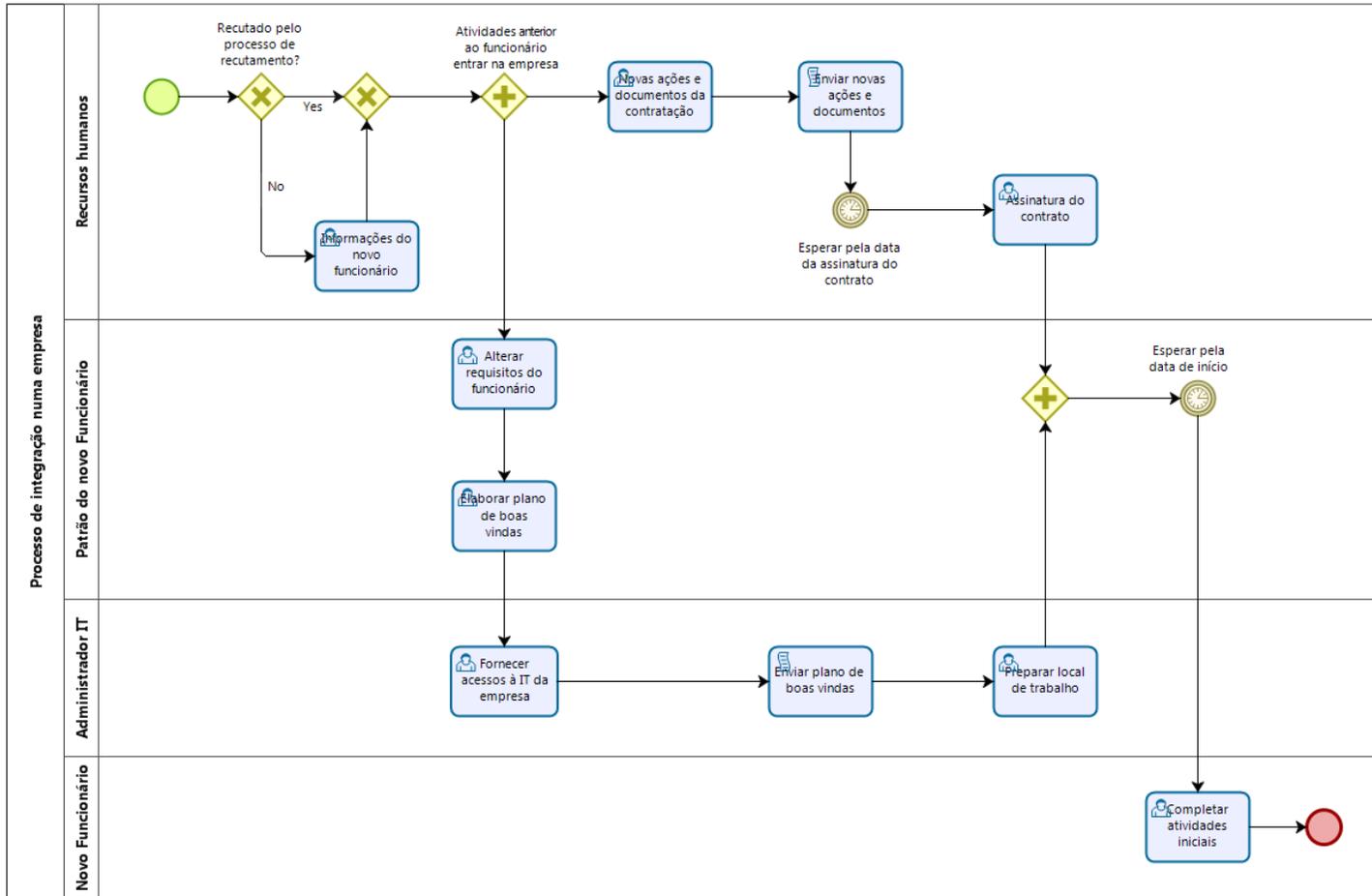


Figura C.7: Modelo fácil do grupo C

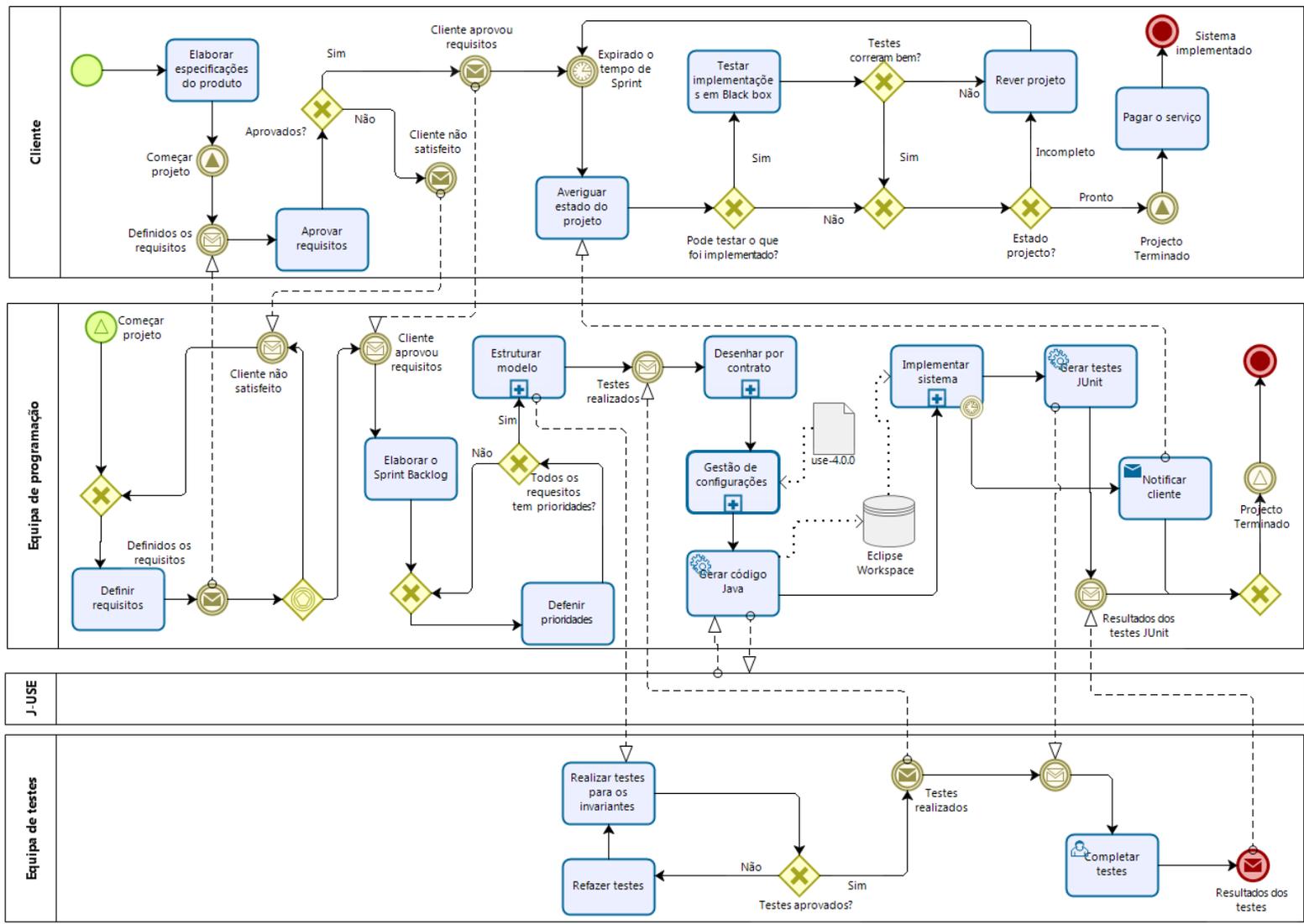


Figura C.9: Modelo difícil do grupo C

Legenda dos modelos

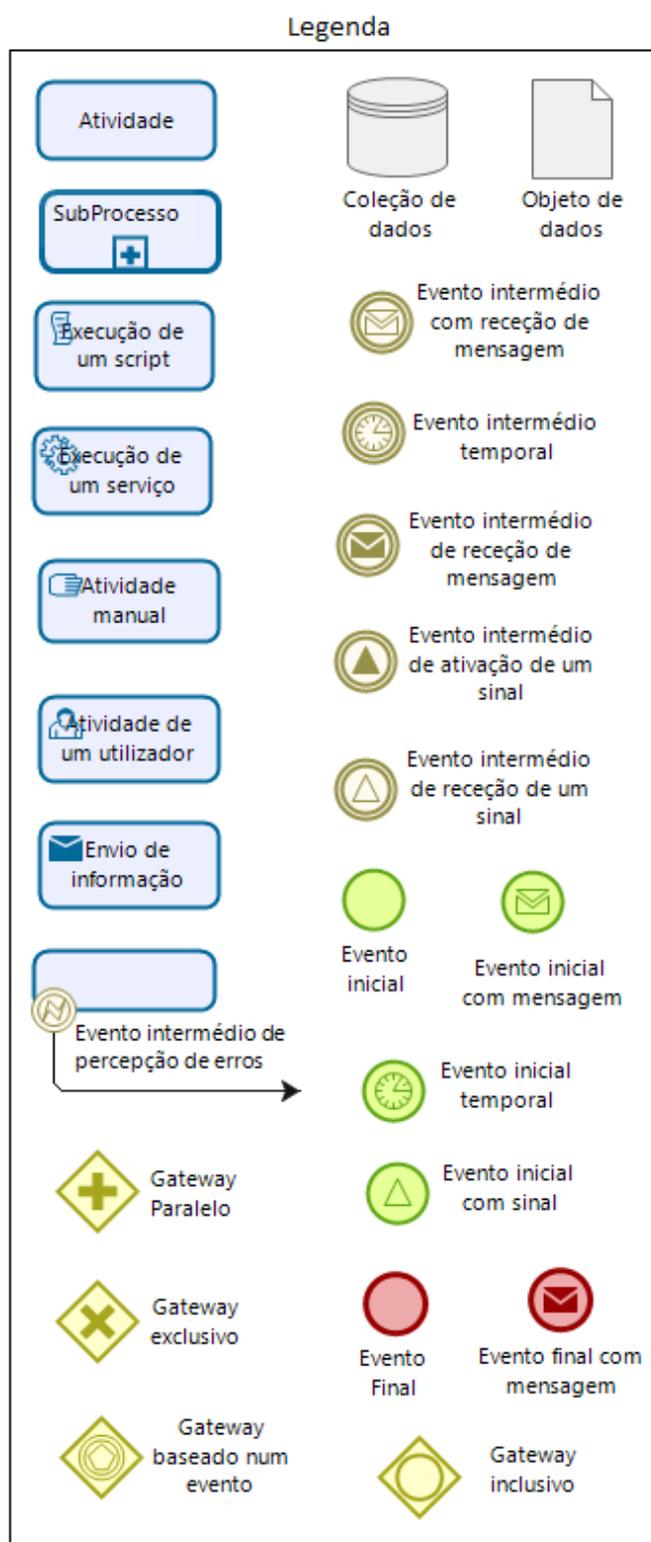


Figura C.10: Legenda dos modelos



QUESTIONÁRIOS DOS TESTES COM VOLUNTÁRIOS

D.1 Grupo A

- Q1 - Quantos participantes existem no modelo?

– Fácil:	– Médio:	– Difícil:
– 0	– 2	– 0
– 1	– 0	– 1
– 2	– 3	– 3
– 3	– 1	– 5
– 4	– 4	– 7

- Q2 - Todos os participantes são essenciais para a conclusão do processo?

- Verdadeiro
- Falso

- Q3 - Quantas atividades são ativadas através de um acontecimento externo ou evento?

– Fácil:	– Médio:	– Difícil:
– 5	– 0	– 0
– 2	– 4	– 2
– 3	– 8	– 4
– 1	– 5	– 6
– 0	– 1	– 8

- Q4 - Quantos participantes utilizam/geram dados para as suas atividades processarem/utilizarem?

- | | | |
|----------|----------|------------|
| – Fácil: | – Médio: | – Difícil: |
| – 5 | – 0 | – 0 |
| – 1 | – 1 | – 1 |
| – 0 | – 2 | – 2 |
| – 3 | – 3 | – 3 |
| – 4 | – 4 | – 4 |

- Q5 - Todos os fluxos de ação divididos voltam a juntar-se?

- Verdadeiro
- Falso

D.2 Grupo B

- Q1 - Quantos sub-processos existem no processo?

- | | | |
|-----------------------------|-----------------------------|-----------------------------|
| – Fácil: | – Médio: | – Difícil: |
| – Não existem sub-processos | – Não existem sub-processos | – Não existem sub-processos |
| – 1 | – 1 | – 3 |
| – 2 | – 2 | – 1 |
| – 3 | – 3 | – 2 |
| – 4 | – 4 | |

- Q2 - Algum dos sub-processos pode ser executado mais que uma vez? Se sim, quantos?

- Não, não existem sub-processos
- Não
- Sim, 3
- Sim, 1
- Sim, 2

- Q3 - Todos os sub-processos são executados obrigatoriamente para a conclusão do processo?

- Não existem sub-processos
- Verdadeiro

- Falso
- Q4 - Algum dos sub-processos recorre a dados para ser executado? Se sim, quantos?
 - Não, não existem sub-processos
 - Não
 - Sim, 3
 - Sim, 1
 - Sim, 2
- Q5 - Quantos sub processos são ativados através de um acontecimento externo ou evento?
 - Não existem sub-processos
 - 0
 - 2
 - 1
 - 3

D.3 Grupo C

- Q1 - Quantos fluxos paralelos existem no modelo?
 - 0
 - 1
 - 2
 - 3
 - 4
- Q2 - Quantos ciclos existem no processo?

- Fácil:	- Médio:	- Difícil:
- Não existem ciclos	- Não existem ciclos	- Não existem ciclos
- 1	- 2	- 1
- 2	- 4	- 2
- 3	- 6	- 3
	- 7	- 4

- A exigência temporal para este teste foi a adequada?
 - Concordo Totalmente
 - Concordo
 - Sem opinião
 - Discordo
 - Discordo Totalmente

- Como considera o seu esforço aplicado durante o teste?
 - Muito Baixa
 - Baixa
 - Normal
 - Elevada
 - Muito Elevada

- Qual o seu nível de frustração durante o teste?
 - Muito Baixa
 - Baixa
 - Normal
 - Elevada
 - Muito Elevada

- Como classificaria a sua performance no teste?
 - Muito Baixa
 - Baixa
 - Normal
 - Elevada
 - Muito Elevada

CARTA DE CONSENTIMENTO

O presente trabalho experimental está a ser realizado no Departamento de Informática da Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa (DI-NOVA), no âmbito de uma tese de mestrado.

Todas as informações declaradas como parte desta experiência são confidenciais e serão mantidas como tal.

O Prof. Miguel Goulão e o Prof. João Araújo são os professores responsáveis pela tese de mestrado no qual este estudo se encontra e por esta experiência. Podem ser contactados através: **mgoul@fct.unl.pt; +351 21 294 85 36 (ext 10731); Escritório: P2/17 e joao.araujo@fct.unl.pt; +351 21 294 85 36 (ext 10747); Escritório: P2/3**

Pode ainda entrar em contacto com a aluna responsável pela tese de mestrado através: **ar.henriques@campus.fct.unl.pt**

Gostaríamos de enfatizar que:

- A sua participação é inteiramente voluntária;
- É livre de se recusar a responder a qualquer pergunta;
- Pode retirar-se da experiência a qualquer momento.

A experiência será mantida estritamente confidencial e será disponibilizada apenas aos professores e aluna da tese de mestrado em questão ou, no caso de avaliação externa de qualidade, avaliadores nas mesmas condições de confidencialidade. Os dados recolhidos nesta experiência podem fazer parte de um relatório de pesquisa final, mas sob nenhuma circunstância nenhum elemento de identificação será incluído no relatório.

Almada, junho de 2018

Ana Henriques

QUESTIONÁRIO DEMOGRÁFICO

As informações pedidas referentes ao participante dos testes são:

- Email
- Idade
- Género
 - M
 - F
- Curso
- Nacionalidade
- Visão
 - Sem acessórios de visão
 - Com óculos
 - Com lentes
- Ano do curso
 - 1º ano da Licenciatura
 - 2º ano da Licenciatura
 - 3º ano da Licenciatura
 - 1º ano do Mestrado

- 2º ano do Mestrado
- Doutorado
- Curso Finalizado

- Ocupação atual
 - Estudante
 - Trabalhador
 - Investigador
 - Trabalhador estudante

- Conhecimento sobre BPMN
 - Sem conhecimento prévio ao presente teste
 - Com conhecimento adquirido em ambiente acadêmico
 - Com conhecimento adquirido em ambiente profissional
 - Com conhecimento adquirido de forma espontânea

Caso a última questão seja uma resposta relacionada com a existência de conhecimento prévio dos modelos, é pedido ainda as seguintes informações:

- Duração de interação
- Altura de interação

MANUAL DE UTILIZADOR



Figura G.1: Página principal da plataforma

G.1 Login na plataforma

Para que seja possível fazer o login no *framework*, deve-se seguir os quatro passos apresentados:

1. Clicar sobre o botão *Login* presente no canto superior direito da página inicial do *framework*.



Figura G.2: Localização da funcionalidade de *Login* da plataforma

2. Após ativar a funcionalidade, é apresentado um *pop-up*, onde se deve preencher com as credenciais de utilizador do *framework*.

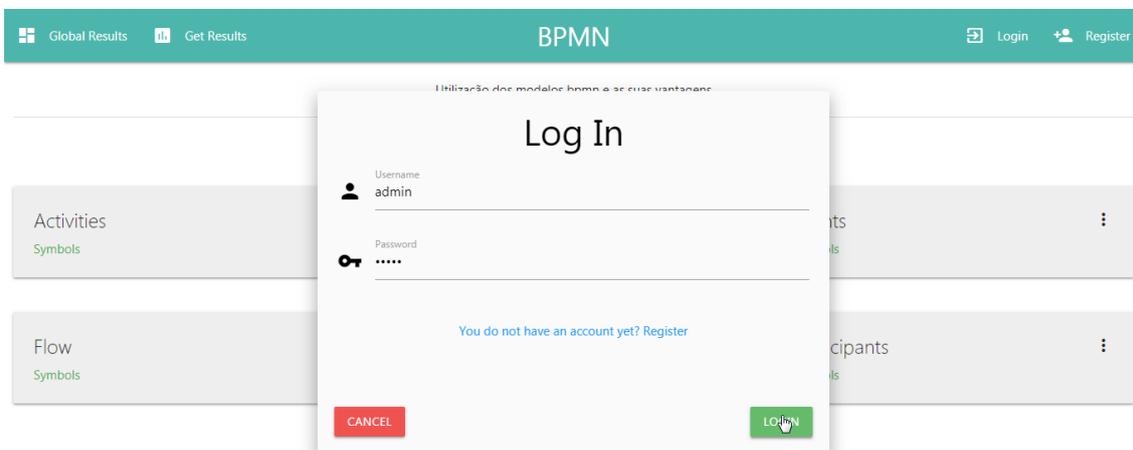


Figura G.3: *Pop-up* de *Login*

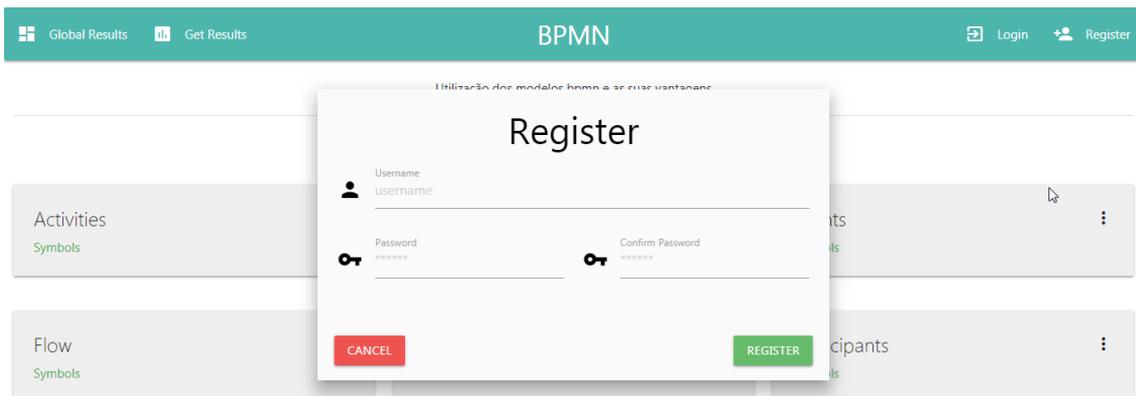
3. Após preencher os dados, deve-se clicar no botão verde *Login* no canto inferior direito do *pop-up*.
4. Depois de feita a verificação dos dados inseridos, o utilizador acede à plataforma.

Figura G.4: Página de utilizador *Logado* na plataforma

G.2 Registo na plataforma

Para que seja possível fazer o registo no *framework*, deve-se seguir os seguintes passos apresentados:

1. Clicar sobre o botão *Register* presente no canto superior direito da página inicial do *framework*.

Figura G.5: *Pop-up* do registo na plataforma

2. Após ativar a funcionalidade, é apresentado um *pop-up*, onde se deve preencher com os dados pedidos para se registar na plataforma
3. Após preencher os dados, deve-se clicar no botão verde *Register* no canto inferior direito do *pop-up*.
4. Depois de proceder ao registo, pode-se então proceder ao *Login* na plataforma.

G.3 Submissão de um modelo no formato XPD

Para que seja possível receber os resultados das métricas de avaliação da complexidade dos modelos [BPMN](#), deve-se seguir os seguintes passos:

1. Em primeiro lugar deve-se clicar na opção *Get Results* disponível no campo superior esquerdo da plataforma.

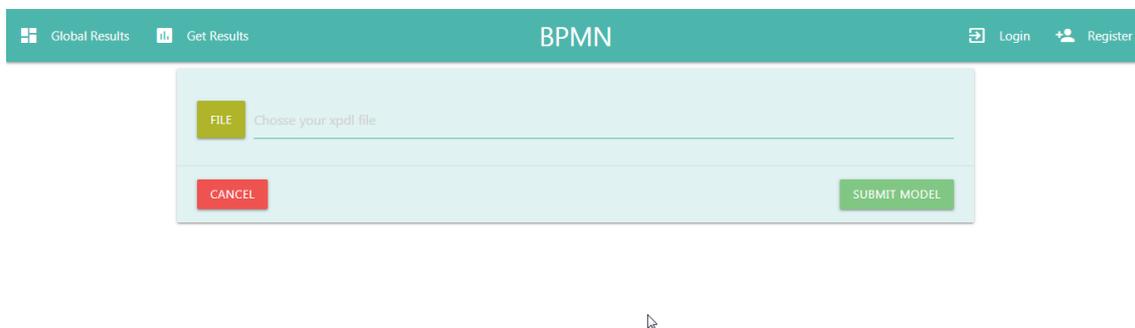
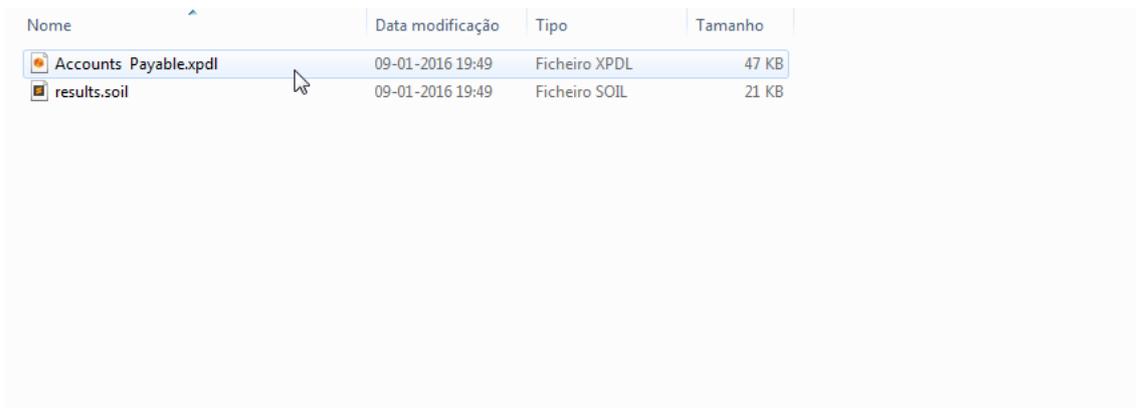


Figura G.6: Página de submissão de modelos [BPMN](#) para análise

2. Após entrar na página, deve-se proceder à escolha do ficheiro pretendido. Tendo em atenção que são apenas analisados ficheiros no formato [XPD](#)

G.3. SUBMISSÃO DE UM MODELO NO FORMATO XPD



Nome	Data modificação	Tipo	Tamanho
Accounts Payable.xpdl	09-01-2016 19:49	Ficheiro XPD	47 KB
results.soil	09-01-2016 19:49	Ficheiro SOIL	21 KB

Figura G.7: Escolha de ficheiros a submeter na plataforma

3. Após a escolha do ficheiro, deve-se clicar no botão verde *Submit Model*.

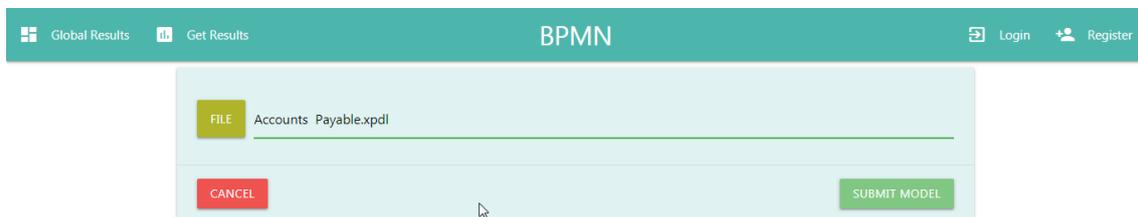


Figura G.8: Apresentação de modelo escolhido para avaliação na plataforma

4. Após a submissão a plataforma irá encaminhar o utilizador para a página onde disponibiliza os resultados das métricas.

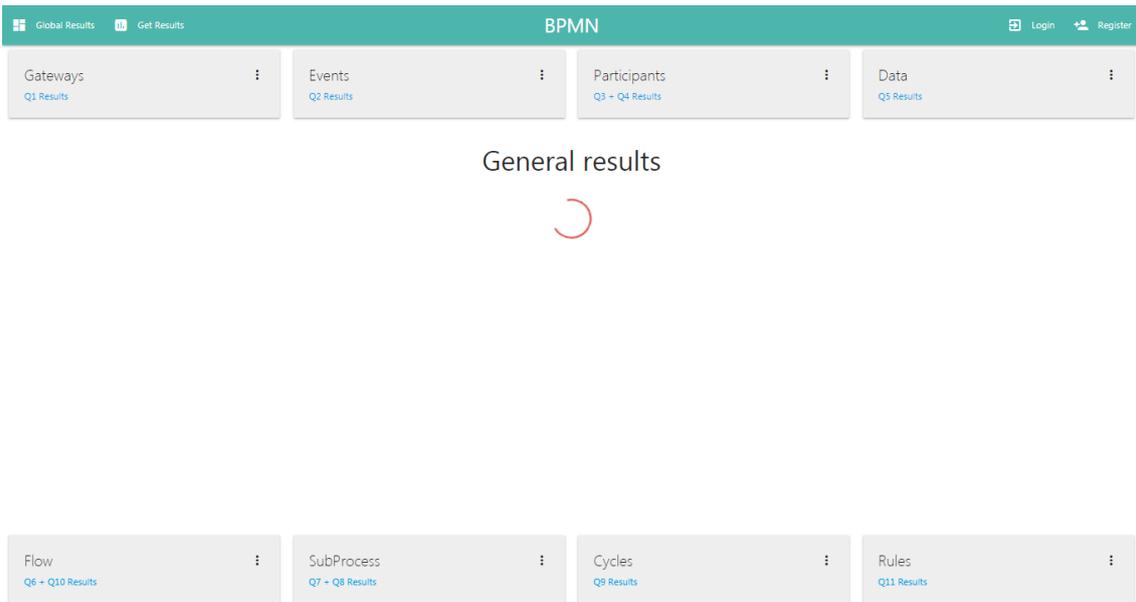


Figura G.9: Página de apresentação dos resultados em *Loading*

5. Depois de serem calculado todos os resultados, são então demonstrados os resultados das métricas, espalhados pela página apresentada.

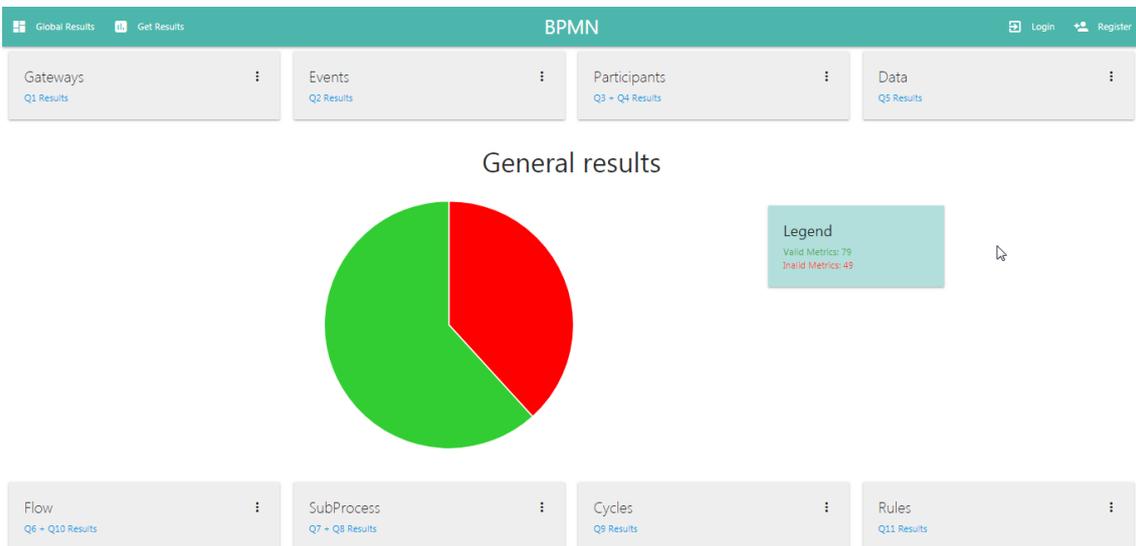


Figura G.10: Página com apresentação dos resultados da aplicação das métricas no modelo submetido

G.4 Realização dos testes para a validação empírica

Para que seja possível receber os resultados das métricas de avaliação da complexidade dos modelos BPMN, deve-se seguir os seguintes passos:

1. Para realizar os testes, o utilizador deve proceder ao *Login* na plataforma e de seguida clicar no botão *Account*.

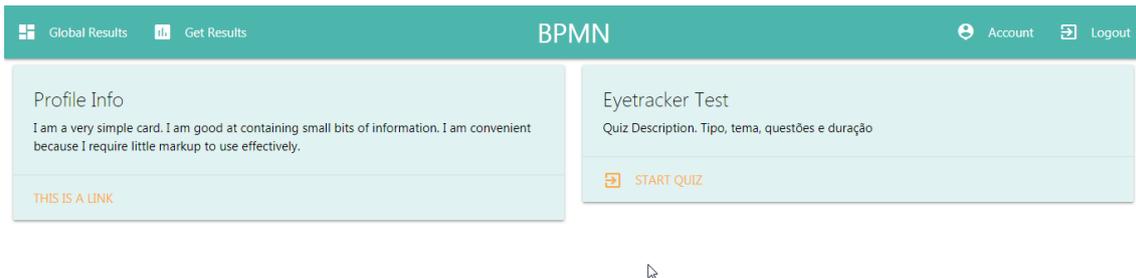


Figura G.11: Página pessoal de um utilizador da plataforma

2. Após entrar na página do perfil do utilizador, deve-se clicar no botão *Start Quiz* e aparecerá o *pop-up* como demonstrado na figura abaixo.

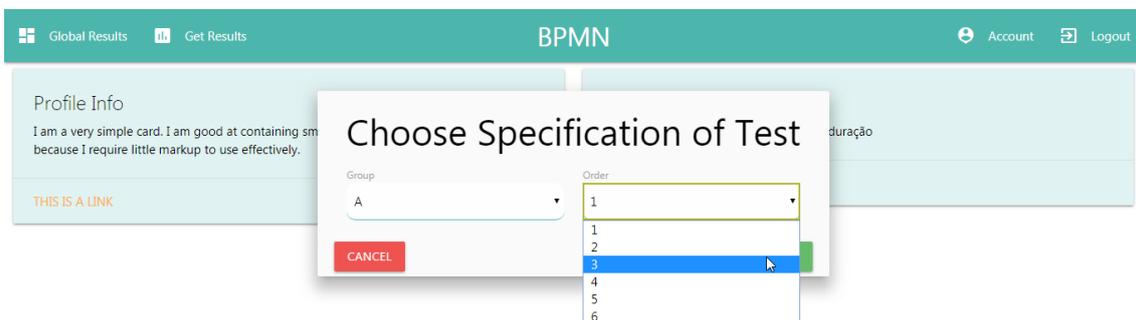


Figura G.12: *Pop-up* de seleção de opções de teste

3. No *pop-up* deve-se escolher qual o grupo e ordem do teste pretendido, clicando de seguida no botão *Start*.
4. Quando o teste inicia, o utilizador deve ler a carta de consentimento e de seguida clicar no botão *Start Test* no canto inferior direito da carta.

Carta de Consentimento

O presente trabalho experimental está a ser realizado no Departamento de Informática da Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa (DI-NOVA), no âmbito de uma tese de mestrado.

Todas as informações declaradas como parte desta experiência são confidenciais e serão mantidas como tal.

O Prof. Miguel Goulão e o Prof. João Araújo são os professores responsáveis pela tese de mestrado no qual este estudo se encontra e por esta experiência. Podem ser contactados através:
mgoul@fct.unl.pt; +351 21 294 85 36 (ext 10731); Escritório: P2/17.
joao.araujo@fct.unl.pt; +351 21 294 85 36 (ext 10747); Escritório: P2/3

Pode ainda entrar em contacto com a aluna responsável pela tese de mestrado através:
ar.henriques@campus.fct.unl.pt

Gostaríamos de enfatizar que:
-A sua participação é inteiramente voluntária;
-É livre de se recusar a responder qualquer pergunta;
- Pode retirar-se da experiência a qualquer momento.

A experiência será mantida estritamente confidencial e será disponibilizada apenas aos professores e aluna da tese de mestrado em questão ou, no caso de avaliação externa de qualidade, avaliadores nas mesmas condições de confidencialidade. Os dados recolhidos nesta experiência podem fazer parte de um relatório de pesquisa final, mas sob nenhuma circunstância nenhum elemento de identificação será incluído no relatório.

Almada, junho de 2018
Ana Henriques

START TEST

Figura G.13: Carta de consentimento de realização e recolha de dados no teste

5. O teste inicia com o primeiro modelo e a sua primeira pergunta, variando a pergunta cinco vezes ao longo do modelo. Aparecendo três modelos ao longo do teste.

G.4. REALIZAÇÃO DOS TESTES PARA A VALIDAÇÃO EMPÍRICA

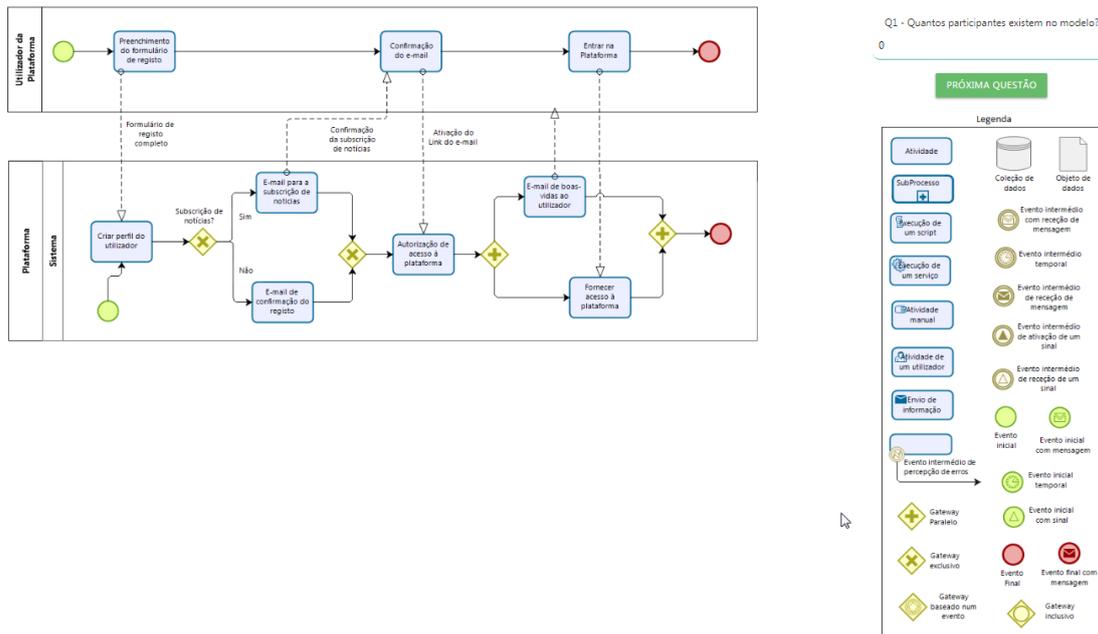


Figura G.14: *Layout* de apresentação de modelo dos testes

- Após a conclusão de cada modelo do teste, o utilizador deve proceder à resposta do teste de usabilidade. Este aparece três vezes ao longo do teste, intercalado com os modelos.

O formulário contém as seguintes perguntas e opções:

- Q1 - A exigência mental para este teste foi a adequada?
 - Concordo Totalmente
- Q2 - A exigência física para este teste foi a adequada?
 - Concordo Totalmente
- Q3 - A exigência temporal para este teste foi a adequada?
 - Concordo Totalmente
- Q4 - Como considera o seu esforço aplicado durante o teste?
 - Muito Baixo
- Q5 - Qual o seu nível de frustração durante o teste?
 - Muito Baixa
- Q6 - Como classificaria a sua performance no teste?
 - Muito Baixa

Botão: **PRÓXIMO MODELO**

Figura G.15: *Layout* de apresentação do teste de usabilidade do teste

7. O passo final do teste consiste em responder a um questionário demográfico, e após concluído, terminar o teste ao clicar no botão *Fim* no canto inferior do teste.

Questionário Demográfico

Email (preencher caso pretenda receber os seus resultados) email@gmail.com		Idade 20	Género ▼
Curso Informática	Nacionalidade Portuguesa	Visão ▼	
Ano do Curso ▼	Ocupação Atual ▼		
Conhecimento sobre BPMN ▼			

FIM

Figura G.16: Questionário demográfico dos testes



META-MODELO BPMN

O esquema na [Figura I.1](#) não apresenta um meta-modelo completo dos BPMN, devido à sua larga semântica e notação. No entanto são apresentadas as ligações e heranças de mais alto nível.

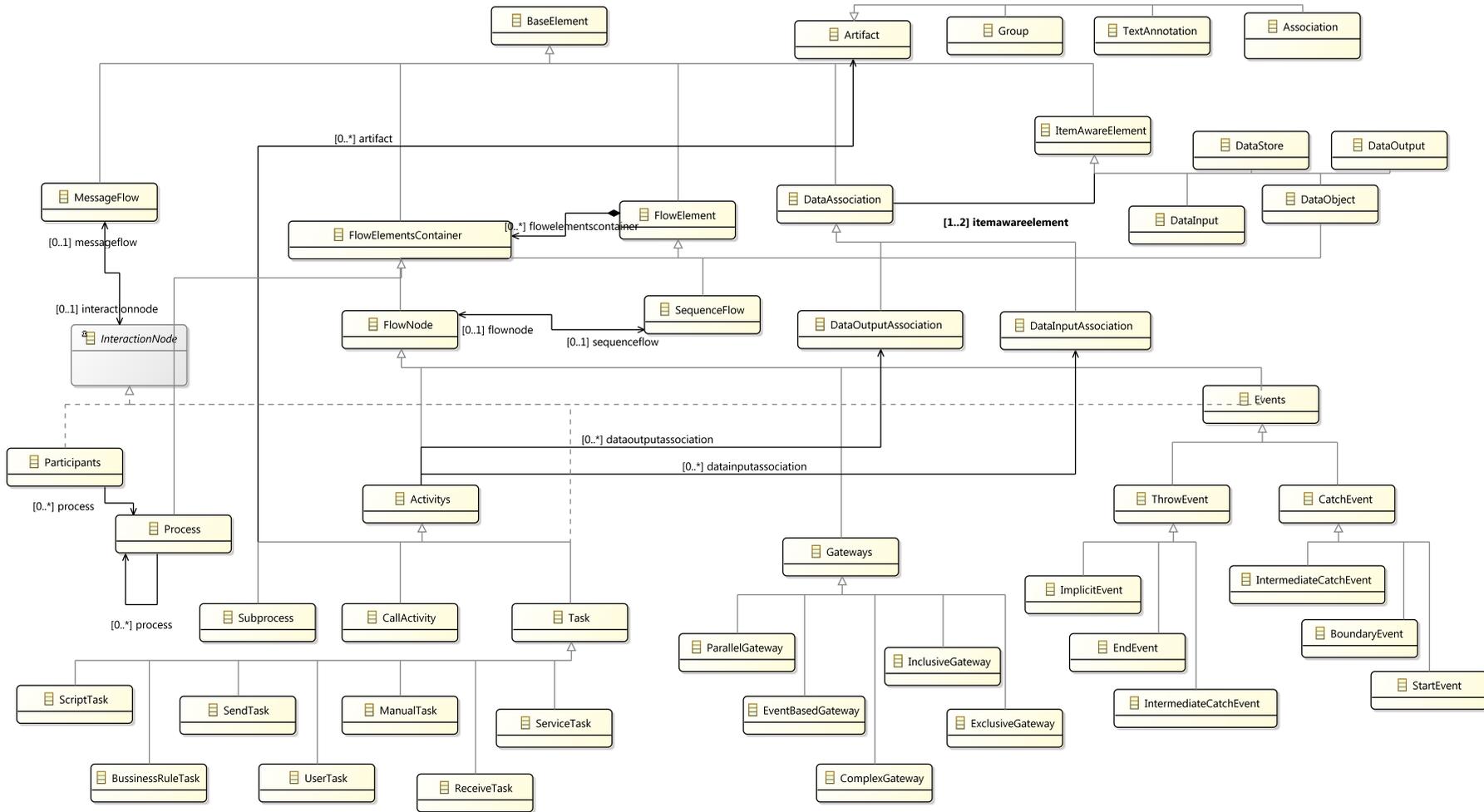


Figura I.1: Meta-modelo BPMN



SIMBOLOGIA MODELOS BPMN

Nesta secção são apresentados os diversos símbolos utilizados para representar os diferentes elementos de um modelo BPMN. Esta simbologia é composta por quatro elementos base, sendo que estes por sua vez, vão conter diversos sub-elementos, cada um com a sua funcionalidade e objetivo [BPMed; GRed].

O primeiro elemento base são os eventos, e estes são compostos por três tipos diferentes de eventos:

- | | | | |
|-------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|
|  | Evento de início. Este ainda pode ser mais seletivos no que toca à iniciação de determinados fluxos ou atividades dos processos: | | |
|  | início de um evento interruptivo com uma mensagem |  | início de um evento não interruptivo com uma mensagem |
|  | início de um evento interruptivo com uma ação temporal |  | início de um evento não interruptivo com uma ação temporal |
|  | início de um evento interruptivo com uma condição |  | início de um evento não interruptivo com uma condição |
|  | início de um evento interruptivo quando é intercetado um determinado sinal |  | início de um evento não interruptivo quando é intercetado um determinado sinal |

	início de múltiplos eventos interruptivos		início de múltiplos eventos não interruptivos
	início de um evento paralelo interruptivo		início de um evento paralelo não interruptivo
	início de um evento interruptivo que ativa um sub-processo		início de um evento não interruptivo que ativa um sub-processo
	início de um evento de erro		início de um evento de compensação
	Evento Intermédio. Sendo que podem existir os seguintes sub-tipos de eventos intermédios:		
	Perceção da receção de uma mensagem interruptiva, estando esta associada a uma atividade ou não		Perceção da receção de uma mensagem não interruptiva
	Envio de uma mensagem para um serviço exterior		Ativação e receção de notificações de um <i>timer</i> interruptivo
	Ativação de um <i>timer</i> não interruptivo		Especificação de um evento que deve ser ativado de uma maneira interruptiva
	Especificação de um evento que deve ser ativado de uma maneira não interruptiva		Mecanismo de espera até a receção de um sinal de uma maneira interruptiva
	Mecanismo de espera até a receção de um sinal de uma maneira não interruptiva		Ativação de um sinal para todos os <i>handlers</i> ativos



Ativação e percepção da recepção de múltiplos eventos interruptivos



Percepção da recepção de múltiplos eventos não interruptivos



Ativação de múltiplos eventos



Ativação de uma atividades em paralelo interruptivas, ou seja, todas tem que ser concluídas antes da próxima fase.



Ativação de uma atividades em paralelo não interruptivas, ou seja, todas tem que ser concluídas antes da próxima fase.



Recepção de múltiplos escalonamentos de uma atividade, ativação de um sub-processo interruptivo



Ativação de um sub-processo não interruptivo



Ativação de um escalonamento através do desencadeamento de uma determinada atividade



Percepção da ocorrência de um erro



Percepção da ativação de um fluxo de compensação



Ativação de um fluxo de compensação



Percepção da ativação de um *link*



Ativação de um *link*



Cancelamento de evento de um sub-processo ou transação



Evento Final. À semelhança com os outros eventos, estes também se podem tornar mais seletivos, recorrendo aos seguintes sub-tipos de eventos finais:



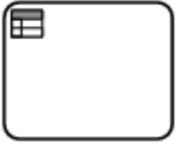
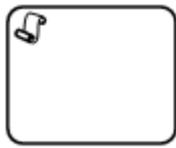
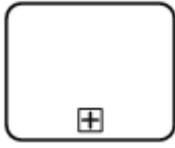
Evento final ativado por uma mensagem

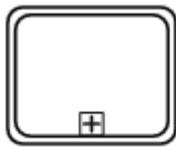


Evento final ativado pelo tempo

	Evento final de vários fluxos		Evento final de um sub-processo
	Evento final invocado devido a um erro		Evento final das tarefas de compensação
	Evento final que cancela um determinado fluxo		Evento final de conclusão de um fluxo ou tarefa

O segundo elemento base são as atividades, sendo que estas por sua vez, têm três subtipos de atividades:

	Tarefa abstrata. Sendo que as tarefas podem ser especificadas consoante a ação em causa:		
	Tarefa para a realização de um serviço		Tarefa para o envio de informação
	Tarefa para a receção de informação		Tarefa realizada por um utilizador
	Tarefa de realização manual		Tarefa para as regras de negócio
	Tarefas de execução de um <i>script</i>		
	Sub-Processo. Estes por sua vez, também podem ser mais seletivos, usando os seguintes tipos de sub-processos:		
	Sub-Processo adhoc (sem conexões com fluxos de sequência)		



Transação entre tarefas



Sub-Processo com eventos, estes sub-processos, podem ser tão seletivos como o tipo de eventos existentes.



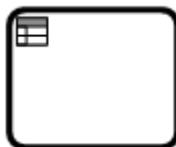
Tarefa de invocação (do inglês *Call Activity*), que também pode ser especifica em:



Invocação de uma tarefa feita por um utilizador



Invocação manual de uma tarefa



Invocação de uma tarefa de regras do negócio



Invocação de tarefas através de um *script*

Os *gateways* podem ser decompostos em oito diferentes tipos:



Gateway Exclusivo sem um indicador



Gateway Exclusivo com um indicador



Gateway Inclusivo



Gateway Paralelo



Gateway Complexo



Gateway baseado num evento



Gateway baseado num evento que desencadeia o início de um processo ou atividade



Gateway Paralelo baseado num evento que desencadeia o início de um processo ou atividade

Por último, o fluxo presente nos diagramas dos modelos BPMN podem representar quatro diferentes tipos de fluxos entre os elementos do modelo [Pol13]:

- Fluxo de sequência
- Fluxo de sequência com condições
- Fluxo de sequência *default*
- Fluxo de mensagem
- Fluxo de uma mensagem de inicialização
- Fluxo de uma mensagem intermédia ou final
- Associação, podendo esta ser com ou sem dados
- Associação com um único sentido
- Associação bidirecional

Para além dos elementos base apresentados, existem ainda a representação dos dados utilizados e gerados ao longo da execução do processo em si, estes podem ser:

- Objeto
- Coleção de objetos
- Dados de Entrada
- Coleção de Dados de Entrada
- Dados de Saída
- Coleção de Dados de Saída
- Armazenamento de dados

Podem existir artefactos nestes modelos, sendo estes representados através de:



Grupo



Anotação em texto

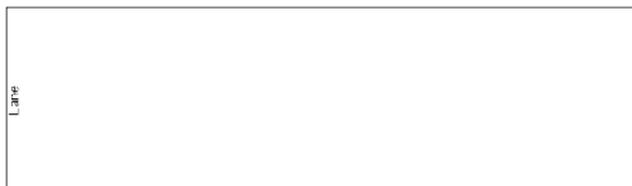
Os participantes destes modelos são representados através de *pools* ou *lanes*. Sendo estas tanto podem ser representadas na vertical como na horizontal. A diferença entre estas duas está relacionada com a quantidade de participantes representados, ou seja, uma *pool*, que representa um processo, pode conter várias *lanes*.



Pool



Pool com instâncias de múltiplos participantes



Lane

A N E X O



MÉTRICAS STANDARD

As seguintes métricas irão avaliar: **Atividades (A)**, **Fluxo de controlo (CF)**, **Fluxo de dados (DF)** e **recursos (R)**.

Na [Tabela III.1](#) são apresentadas as métricas gráficas de acordo com o seu domínio. A grande maioria destas métricas são retiradas da topologia dos modelos [[HVM11](#)].

	Métrica	Componentes avaliadas	Fonte
ES	4. Complexidade do fluxo de controlo	CF	[Car05a]
	44. Contagem de nós	CF	[GL07]
	66. Indicador da complexidade dos <i>Gateways</i>	CF	[SG+12b]
BPM	56. Complexidade Estrutural de <i>Cheng</i>	A, CF, DF	[Ço14]
	59. Complexidade da usabilidade de <i>Cheng</i>	DF, R	[Che08]
	64. Complexidade do modelo de <i>Cheng e Prabhu</i>	A, CF	[CP08]
Sem definição	7. Coeficiente da complexidade da rede	A, CF	[LK01]
	11. Diâmetro	CF	[SG+12a]
	12. Densidade	CF	[SG+12a]
	13. Média de <i>Gateways</i> /grau de conexões	CF	[SG+12a]
	14. Máximo de <i>Gateways</i> /grau de conexões	CF	[SG+12a]
	16. <i>Sequentiality</i>	CF	[SG+12a]
	18. Separabilidade	A, CF	[SG+12a]
	20. <i>Cyclicity</i>	A, CF	[SG+12a]
	39. Índice de Complexidade	A, CF	[LK01]
	40. Estimativa de restrições	A	[LK01]
	43. Anti-Padrões	A, CF	[GL07]
	46. Média dos graus dos vértices	CF	[Men]

Tabela III.1: Métricas de gráficos de acordo com o seu domínio

Na Tabela III.2 são apresentadas as métricas relacionadas com a notação gráfica dos modelos de acordo com o seu domínio.

	Graphical Notation Definition	Componentes avaliadas	Fonte
ES	66. Indicador da complexidade dos <i>gateways</i>	CF	[SG+12b]
BPM	19. Paralelismo	CF	[SG+12a]
	21. Número total de fluxos de sequência	CF	[SG+12a]
	22. Número total de eventos	CF	[SG+12a]
	23. Número total de <i>Gateways</i>	CF	[SG+12a]
	24. Número de fluxos de sequência de eventos	CF	[SG+12a]
	25. Número de fluxos de mensagens	DF	[SG+12a]
	26. Número de fluxos de sequência de <i>Gateways</i>	CF	[SG+12a]
	27. Nível de conectividade entre <i>pools</i>	CF	[SG+12a]
	28. Número total de objetos de dados	DF	[SG+12a]
	29. Número de decisões inclusivas	CF	[SG+12a]
	30. Número de bifurcações paralelas	CF	[SG+12a]
	31. Número de <i>pools</i>	CF	[SG+12a]
	32. Número de decisões complexas	CF	[SG+12a]
	33. Número de decisões exclusivas dos <i>gateways</i> baseadas nos dados	CF, DF	[SG+12a]
	34. Número de decisões exclusivas dos <i>gateways</i> baseadas nos eventos	CF	[SG+12a]
	35. Número de eventos de mensagens intermediários	DF	[Rol+06]
	36. Número de eventos de mensagens finais	DF	[Rol+06]
	37. Número total de de processos colapsáveis	A	[Rol+06]
	38. Nível de conectividade entre atividades	A, CF	[Rol+06]
	47. Quantidade de decisões por <i>pool</i>	CF, R	[Deb+10]
	48. Quantidade de tarefas executadas por <i>pool</i>	A, R	[Deb+10]
	49. Quantidade de tarefas nas secções das <i>pools</i>	A, R	[Deb+10]
	50. Proporção da distribuição de tarefas pelos participantes	A, R	[Deb+10]
	51. Proporção de tarefas pelas secções de uma <i>pool</i>	A, R	[Deb+10]
	52. Número de sub-processos por <i>pool</i>	A, R	[Deb+10]
	53. Quantidade de fluxos de mensagens entre duas <i>pools</i>	DF	[Deb+10]

Tabela III.2: Métricas sobre a notação gráfica de acordo com o seu domínio

Na Tabela III.3 são apresentadas as métricas cuja teoria é classificada como *Petri nets*. Estas recorrem a métodos matemáticos e gráficos para serem calculadas. Os métodos gráficos conseguem extrair as informações visuais dos modelos e os métodos matemáticos conseguem extrair informações através da aplicação de formulas e equações matemáticas [Mur89].

	Petri Nets	Componentes avaliadas	Fonte
ES	15. Número de <i>handles</i> ou <i>Gateways</i> incompatíveis	CF	[SG+12a]
	41. Métrica estendida de Cardoso	CF	[LA09]
	42. Métrica cíclica estendida	A, CF	[LA09]
	66. Indicador da complexidade dos <i>Gateways</i>	CF	[SG+12b]
Sem defenição	8. Rácio entre as divisões e junções de fluxos	CF	[GL06b]
	62. Métrica de estruturação	A, CF	[SG+12b]

Tabela III.3: Métricas Petri Nets de acordo com o seu domínio

Na Tabela III.4 são apresentadas as métricas da Complexidade do *Software*. Estas tratam de medir a interação entre um *software* e a sua execução [Kea+86].

	Software Complexity	Componentes avaliadas	Fonte
ES	1. Número de atividades/tarefas	A	[Car+06]
	2. Número de atividades e fluxos de controlo	A, CF	[Car+06]
	3. Número de atividades, junções e divisões	A, CF	[Car+06]
	5. Complexidade do processo baseada em <i>Halstead</i>	A, CF, DF	[Car+06]
	6. Complexidade da interface (<i>fan-out</i> e <i>fan-in</i>)	DF	[Car+06]
	9. <i>Nesting depth</i>	CF	[GL06b]
	45. Complexidade logarítmica	A, CF	[Car07]
BPM	61. Métricas de complexidade baseadas em <i>GQM</i>	A, CF	[Muk08]
	65. Antonini	A, CF, DF, R	[Ant+11]

Tabela III.4: Métricas da complexidade de *software* de acordo com o seu domínio

Na Tabela III.5 são apresentadas as métricas do nível cognitivo. Tal como o nome indica, estas pretendem medir a capacidade dos utilizadores realizarem as suas tarefas no produto que é desenvolvido.

	Cognitive Load	Componentes avaliadas	Fonte
ES	10. Complexidade cognitiva	A, CF	[GL06a]
BPM	56. Profundidade dos arcos de fluxo de controlo das atividades	A, CF, DF	[Ço14]
	58. Complexidade de Interação de <i>Cheng</i>	DF	[Che08]
	59. Complexidade da usabilidade de <i>Cheng</i>	DF, R	[Che08]
	63. Conectividade cruzada	CF	[Van+08]

Tabela III.5: Métricas da capacidade cognitiva de acordo com o seu domínio

Na [Tabela III.6](#) são apresentadas as métricas classificadas como *Shannon's Information Entropy*. Estas propõem-se a medir a informação existente de acordo com a distribuição dos diferentes elementos no modelo [GV04].

	Shannon's Information Entropy	Componentes avaliadas	Fonte
BPM	57. Complexidade Estrutural de <i>Cheng</i>	A, CF	[Che08]
	58. Complexidade de Interação de <i>Cheng</i>	DF	[Che08]
	60. Complexidade operacional total de <i>Cheng</i>	A, CF, DF, R	[Che08]

Tabela III.6: Métricas de Shannon de acordo com o seu domínio

Na [Tabela III.7](#) são apresentadas as métricas numéricas.

	Number	Componentes avaliadas	Fonte
Sem defenição	54. Métrica quadrática de <i>Durfee</i>	A, CF	[KN12]
	55. Métrica quadrática perfeita	A, CF	[KN12]

Tabela III.7: Métricas numéricas de acordo com o seu domínio

A N E X O
IV

MÉTRICAS STANDARD EYE-TRACKER

Nas seguintes tabelas são apresentadas métricas para a análise de dados referentes à: **Número de fixações, Duração das fixações, Movimentos sacadas** e por fim os **Movimentos scanpath**.

Nome	Fórmula	Fonte
Número de Fixações	FC = Número total de fixações AOI	[CS90; Cro+02; Sha+13; SM10; Sha+12; Tur+14; Uwa+06; Yus+07]
Taxa de Fixações	$RF = \frac{\text{Número total de fixações AOI}}{\text{Número total de fixações AOG}}$	[Bin+13; DS+14; PG10; SM10; Sha+12; Tur+14]
Densidade Espacial	$SD = \frac{\sum_{i=1}^n c_i}{\text{Número de fixações numa dada área}}$	[DS+14; Soh+12]
Área convexa dos extremos	Área convexa mais pequena do conjunto de fixações	[Sha+13; Sha+12; Soh+12]

Tabela IV.1: Métricas relacionadas com o número de fixações

Nome	Fórmula	Fonte
Número de movimentos sacadas	QF = Número total de movimentos sacadas	[Fri+14]
Duração dos movimentos sacadas	QF = Duração total dos moviemtnos sacadas	[Fri+14]

Tabela IV.2: Métricas relacionadas com os movimentos sacadas

Nome	Fórmula	Fonte
Duração média das fixações	$AEF(AOI) = \frac{\sum_{i=1}^n (ET(F_i) - ST(F_i) dos AOI)}{n}$	[Bin+13; CS90; PG10; Sha+13; SM10; Soh+12]
Taxa do tempo de fixação dos target AOI	$ROAFT = \frac{\sum_{i=1}^n (ET(F_i) - ST(F_i) dos AOI)}{\sum_{i=1}^n (ET(F_i) - ST(F_i) dos AOG)}$	[Bed12; BT06; Bin+13; PG10; Sha+12]
Tempo de fixação	Duração total dos movimentos AOI	[Bed12; Bus+14; CS90; Cro+02; PM13; Rod+14; Uwa+06]
Duração média das fixações relevantes	$ADRF = \frac{\text{Duração das fixações AOI relevantes}}{\text{Número total de AOI relevantes}}$	[DS+14; Jea+09]
Taxa normalizada das fixações relevantes	$NRRF = \frac{ADRF}{\frac{\text{Duração total de AOI}}{\text{Número total de AOI}}}$	[DS+14; Jea+09; Soh+12]

Tabela IV.3: Métricas relacionadas com a duração das fixações

Nome	Fórmula	Fonte
Matriz transitória	$TM = \frac{\sum_{i=1}^n \sum_{j=1}^n c_{ij}}{\text{Número de fixações numa dada área}}$	[Soh+12]
Frequência da mudança de atenção	Número de trocas entre AOIs	[AC06; BT08]
Posição das <i>scanpath</i>	$\frac{SP \cap RR}{SP}$	[PM13]
<i>Scanpath</i> de recordação	$\frac{SP \cap RR}{RR}$	[PM13]
Medida F de <i>scanpath</i>	$2 * \frac{SPP * SPR}{SSP * SPR}$	[PM13]

Tabela IV.4: Métricas relacionadas com os movimentos *scanpath*



MÉTRICAS BOAS PRÁTICAS DE MODELAÇÃO

De seguida são apresentadas métricas de boa modelação dos modelos *BPMN*, definidas em [Cor14].

Para os processos recomenda-se que:

1. Cada diagrama tenha entre dois a sete *Flow Objects*.

Em relação aos eventos deve-se ter em conta:

1. Não se deve ter tarefas de envio ou receção ou eventos intermediários de mensagens em simultâneo.
2. Deve-se usar eventos de início e final explícitos.
3. Deve-se usar preferencialmente um único evento de início.
4. Usar condições *default* quando se usa fluxos de sequência condicionais.
5. Quando se usa um *gateway* baseado em eventos deve-se usar um evento de tempo intermediário.
6. Um evento deve ter no máximo um fluxo de sequência de saída.
7. O evento de início deve possuir um nome.
8. O evento inicial de mensagem deve possuir um fluxo de mensagem de entrada.
9. Um evento intermediário de receção de mensagem deve ter um fluxo de mensagem de entrada.
10. Um evento intermediário de envio deve ter um fluxo de mensagem de saída.
11. Os eventos intermediários devem possuir um nome.

12. Os eventos finais devem possuir o nome do estado final do processo.
13. Se um sub-processo for seguido de um *gateway* com as hipóteses sim ou não, pelo menos um dos eventos finais do sub-processo deve ter o mesmo que concordância com a *label* do *gateway*.

Em relação os *gateways*, deve-se ter em atenção que:

1. Deve ser usada uma condição *default* nos *gateways* exclusivos.
2. Deve ser usada uma condição *default* nos *gateways* inclusivos.
3. Deve-se fazer coincidir os fluxos de controlo divididos e agregados nos *gateways* paralelos, caso seja pretendido agregar os fluxos.
4. Os *gateways* paralelos devem possuir os fluxos de sequência de entrada corretos.
5. Deve-se fazer coincidir os *gateways* inclusivos de divisão e agregação.
6. Deve-se utilizar um *gateway* para agregar fluxos exclusivos.
7. Deve-se evitar usar *gateways* de agregação e divisão em simultâneo.
8. Deve existir unicamente um fluxo de sequência de saída sem nome nos *gateways* baseados em eventos e exclusivos.
9. Um *gateway* inclusivo deve ter todos os fluxos de sequência de saída identificados.
10. Se um sub processo é seguido de um *gateway* com as hipóteses sim ou não, deve de existir pelo menos um evento final no sub-processo cujo nome coincide com o nome do *gateway*.

Para as atividades, deve-se ter em atenção:

1. As atividades devem possuir um nome.
2. Num mesmo processo não deve de haver repetição no nome das atividades.
3. Uma tarefa de envio deve ter um fluxo de mensagem de saída.
4. Uma tarefa de receção deve ter um fluxo de mensagem de entrada.
5. Se um sub-processo é seguido de um *gateway* com uma pergunta, este deve ter mais que um evento final.

Finalmente, para os fluxos de mensagens deve-se ter em conta que:

1. Os fluxos de mensagem devem ter o mesmo nome que a mensagem em si.