



UHASSELT

KNOWLEDGE IN ACTION

Faculteit Bedrijfseconomische Wetenschappen

master handelsingenieur in de beleidsinformatica

Masterthesis

Onderzoek naar de begripbaarheid van BPMN-modellen

Elke Teuwen

Scriptie ingediend tot het behalen van de graad van master handelsingenieur in de beleidsinformatica

PROMOTOR :

dr. Frank VANHOENSHOVEN



UHASSELT

KNOWLEDGE IN ACTION

www.uhasselt.be
Universiteit Hasselt
Campus Hasselt:
Martelarenlaan 42 | 3500 Hasselt
Campus Diepenbeek:
Agoralaan Gebouw D | 3590 Diepenbeek

2021
2022



Faculteit Bedrijfseconomische Wetenschappen

master handelsingenieur in de beleidsinformatica

Masterthesis

Onderzoek naar de begripbaarheid van BPMN-modellen

Elke Teuwen

Scriptie ingediend tot het behalen van de graad van master handelsingenieur in de beleidsinformatica

PROMOTOR :

dr. Frank VANHOENSHOVEN

Onderzoek naar de begrijpbaarheid van BPMN-modellen

Elke Teuwen,^{1,2}

Promotor: dr. Frank Vanhoenshoven²

¹ Handelsingenieur in de Beleidsinformatica

² UHasselt, Faculteit Bedrijfseconomische Wetenschappen, Agoralaan, 3590 Diepenbeek, België

Samenvatting Business proces modellen worden gebruikt door verschillende stakeholders [7]. Niet alle deze stakeholders hebben theoretische kennis van of praktische ervaring met procesmodellen. De algemene begrijpbaarheid van procesmodellen is daarom van uiterst belang. Deze begrijpbaarheid hangt af van veel verschillende factoren [23, 24, 36]. In deze paper werden, a.d.h.v. een empirisch onderzoek, enkele van deze factoren bestudeerd. Het doel was om te onderzoeken hoe verschillende representaties van modulariteit, condities en activiteit-namen de begrijpbaarheid van procesmodellen in BPMN beïnvloeden. Om dit te onderzoeken werden er voor elke factor meerdere procesmodellen ontworpen. In totaal namen er 66 participanten deel aan het onderzoek. Alle participanten kregen vragen over vier verschillende procesmodellen. Dit resulteerde in een totaal van 264 observaties. De resultaten tonen aan dat bij het toepassen van *modulariteit*, een procesmodel met *expanded sub-processen* makkelijker begrijpbaar is dan een procesmodel met *collapsed sub-processen*. Bij het weergeven van *condities* is er geen verschil in begrijpbaarheid tussen een procesmodel met *XOR-gateways* of met *intermediate error events*. Bovendien werd er in dit onderzoek ontdekt dat het gebruik van *onlogische activiteit-namen* geen impact heeft op de begrijpbaarheid van procesmodellen. Het gebruik van *logische activiteit-namen* heeft echter wel voordelen voor de begrijpbaarheid t.o.v. het gebruik van *abstracte activiteit-namen*. Als laatste werd als extra element in dit onderzoek ook bevestigd dat de kennis van de participanten een grote impact heeft op de begrijpbaarheid van proces modellen.

Keywords: Proces modelleren · Begrijpbaarheid · BPMN · Modulariteit · Condities · Activiteit-namen

1 Introductie

Iedere onderneming bevat bedrijfsprocessen die doorheen hun levenscyclus verschillende stakeholders hebben [7]. Business analisten houden zich bezig met de creatie en verbetering van deze processen [7]. De implementatie en automatisatie wordt meestal uitgevoerd door technologische ontwikkelaars [7]. Bovendien

zorgen de managers en werknemers van een bedrijf voor het beheren en het uitvoeren van de bedrijfsprocessen [7]. Door de verschillende stakeholders is er nood aan een eenduidige manier om bedrijfsprocessen te representeren [4]. Voor deze representatie kunnen verschillende grafische notaties, zoals Petri Nets, BPMN, Flow Charts, Data Flow Diagrams (DFD), etc. gebruikt worden om de bedrijfsprocessen weer te geven via modellen. De momentele leider en de-facto standaard van deze notaties is Business Process Modeling Notation, ook wel BPMN genoemd [13, 28, 40]. BPMN is een grafische notatie waarin duidelijk wordt weergegeven welke taken er worden verricht, door wie deze taken worden verricht en wanneer deze taken precies worden verricht. Buiten het kiezen van de grafische notatie is het ook heel belangrijk om rekening te houden met de begripbaarheid van de gemaakte procesmodellen [15, 41]. Niet alle stakeholders hebben namelijk theoretische kennis van of praktische ervaring met procesmodellen. Het bestuderen van de begripbaarheid van deze modellen is daarom van uiterst belang [2].

De begripbaarheid van procesmodellen kan afhangen van verschillende factoren. Een van de belangrijkste en meest gekende factoren is de grootte van deze modellen [7, 23]. De grootte van een procesmodel heeft een beduidende impact op de begripbaarheid: hoe groter een model, hoe complexer het model meestal is [4, 13, 22–24, 36, 42]. Om de grootte van een procesweergave te verkleinen en te zorgen voor een algemene betere begripbaarheid van procesmodellen, wordt er vaak gebruik gemaakt van de *modulariteit* van procesmodellen om grote modellen op te splitsten in meerdere kleinere modellen [6, 22, 33, 35]. Deze opsplitsing gebeurt meestal door middel van *sub-processen* en wordt verondersteld een positief effect te hebben op de complexiteit en de begripbaarheid van procesmodellen [4, 6, 27, 33, 35, 41]. Tot op het heden zijn er echter slechts een beperkt aantal onderzoeken uitgevoerd die dit daadwerkelijk aantonen. In deze paper wordt hier dieper op ingegaan door de begripbaarheid van *expanded sub-processen* en *collapsed sub-processen* in BPMN-modellen met elkaar te vergelijken.

Een tweede belangrijke factor omvat het modelleren van *condities* en *beslissingen*. Onderzoek toonde o.a. aan dat het gebruik van veel *XOR-gateways*, die zorgen voor splitsingen van de flows in een BPMN-model, een negatieve impact heeft op de begripbaarheid [39]. Om het aantal *XOR-gateways* te verminderen kunnen condities en beslissingen ook weergegeven worden door *intermediate error events*. Als tweede factor bestudeert deze paper de begripbaarheid van *XOR-gateways* en *intermediate error events* en worden deze twee modeleringsopties met elkaar vergeleken.

Een derde belangrijk factor van de begripbaarheid van procesmodellen is het gebruik van *activiteit-namen*. Een voorgaand onderzoek toonde reeds aan dat *abstracte activiteit-namen* een positief effect hebben op de begripbaarheid van BPMN-modellen [25]. In deze paper wordt er dieper ingegaan op de impact van *activiteit-namen* door de begripbaarheid van BPMN-modellen met *abstracte activiteit-namen*, *tekstueel logische activiteit-namen* en *tekstueel onlogische activiteit-namen* te bestuderen.

In de volgende sectie worden enkele concepten rond de begripbaarheid van BPMN-modellen kort toegelicht. Bovendien wordt de gerelateerde literatuur

rond de 3 factoren van de begrijpbaarheid, die het onderwerp van dit onderzoek vormen, kort besproken. Vervolgens wordt in sectie 3 de probleemstelling en de toegevoegde waarde van dit onderzoek omschreven. Sectie 4 omschrijft de opzet en de uitvoering van het empirische onderzoek. In sectie 5 worden de resultaten van het onderzoek uiteengezet en besproken. Vervolgens worden in sectie 6 de beperkingen van dit onderzoek kort aangehaald. Als laatste wordt er in sectie 7 afgesloten met een conclusie waarin de bevindingen van dit onderzoek worden samengevat.

2 Gerelateerde literatuur

In deze sectie wordt dieper ingegaan op de reeds bestaande literatuur rond de begrijpbaarheid van procesmodellen. Allereerst wordt in sectie 2.1 de betekenis van begrijpbaarheid kort toegelicht en worden er enkele methoden aangehaald om deze begrijpbaarheid te operationaliseren. Vervolgens worden in sectie 2.2 de verschillende soorten factoren, die invloed hebben op de begrijpbaarheid van procesmodellen, uiteengezet. In de drie daaropvolgende secties wordt ten slotte dieper ingegaan op drie van deze factoren, die het onderwerp vormen van dit onderzoek. Sectie 2.3 gaat over de *modulariteit* van BPMN-modellen en de impact hiervan op de begrijpbaarheid van deze modellen. Sectie 2.4 gaat over de begrijpbaarheid van *intermediate error events* en *XOR-gateways*. De laatste sectie, sectie 2.5, gaat over de invloed van *activiteit-namen* op de begrijpbaarheid van BPMN-modellen.

2.1 De begrijpbaarheid van procesmodellen

De begrijpbaarheid van een procesmodel wordt gedefinieerd als de mate waarin de informatie, die een procesmodel bevat, makkelijk te begrijpen is door de lezer van dat procesmodel [36]. In de literatuur worden, o.a. via het gebruik van enquêtes, verschillende indicatoren gebruikt om deze begrijpbaarheid te operationaliseren [6]. De meest gebruikte en voor de hand liggende methode is het berekenen van de *understandability task effectiveness* [6, 40, 41]. Dit wordt gedaan door het aantal correct beantwoorde vragen te delen door het totaal aantal vragen [6]. Een andere methode omvat het meten van de *understandability task efficiency* [6]. In sommige onderzoeken gebeurt dit door specifiek het aantal correct beantwoorde vragen te delen door de totale gebruikte tijd [6, 16, 40, 41]. In andere onderzoeken wordt hiervoor het totaal aantal beantwoorde vragen gedeeld door de totale gebruikte tijd [6]. Buiten deze twee objectieve indicatoren zijn er ook nog vier subjectieve indicatoren die gebruikt worden om de begrijpbaarheid van procesmodellen te operationaliseren. In sommige onderzoeken wordt er gevraagd naar de *mentale inspanning* bij het beantwoorden van de vragen [6, 40, 41, 45]. In andere onderzoeken worden er vragen gesteld over het *gebruiksgemak* of wordt er gevraagd naar het *waargenomen nut* van het procesmodel, om op deze manier de begrijpbaarheid af te leiden [6, 40, 41]. Het *waargenomen nut* wordt namelijk deels bepaald door het waargenomen gemak

waarmee modellen worden begrepen [17]. Als laatste wordt er ook soms gevraagd naar de mate waarin de participanten van plan zijn het procesmodel te gebruiken, ook wel de *gebruikersintentie* genoemd [6]. Ook vanuit deze indicator kan de begripbaarheid afgeleid worden [6].

2.2 Factoren van de begripbaarheid

De begripbaarheid van procesmodellen kan afhangen van verschillende factoren die volgens de literatuur onderverdeeld kunnen worden in zes hoofdgroepen: de model gerelateerde factoren, de persoonlijke factoren, de tekstuele inhoud, het modeldoel, de modelleringsnotatie en de visuele lay-out [23, 24, 36].

In de meeste voorgaande studies ligt de focus hoofdzakelijk op de *model gerelateerde factoren*. Het zijn dit soort factoren waar de modelontwerper de meeste controle over heeft. Een van die model gerelateerde factoren is bv. de mate waarin de verschillende taken in een procesmodel met elkaar verbonden zijn, ook wel de cross-connectiviteit genoemd [36]. Maar ook simpelweg het aantal verbindingen in het procesmodel of het aantal AND-joins zijn proces gerelateerde factoren die een impact kunnen hebben op de begripbaarheid [36]. De meest besproken en voor de hand liggende model gerelateerde factor is de grootte van procesmodellen. Verschillende papers bespreken de negatieve impact van groeiende procesmodellen op de begripbaarheid van zulke modellen [4, 13, 23, 24, 36, 42]. Grote procesmodellen zorgen er vaak voor dat de lezer moet scrollen doorheen het procesmodellen en zijn aandacht moet verdelen over meerdere weergaven, dit heeft een negatieve impact op de begripbaarheid [16, 33, 40, 41].

De tweede groep bevat de *persoonlijke factoren*. Onder andere de praktische ervaring met en de theoretische kennis van het modelleren van procesmodellen zijn hier onderdeel van [36]. Deze factoren hebben een grote impact op de begripbaarheid [36]. Volgens een voorgaand onderzoek van Reijers et al. is de impact van deze factoren zelfs groter dan de impact van model gerelateerde factoren [36]. Mendling et al. ondervonden bovendien dat modelleringservaring slechts een kleine impact heeft op de begripbaarheid van procesmodellen [25]. Theoretische kennis, daarentegen, heeft volgens zijn onderzoek een grote impact op de begripbaarheid [25]. Er is dus zeker nood aan theoretische training bij het gebruik van BPMN-modellen [24].

Factoren rond de *tekstuele inhoud* van een procesmodel, zoals de labels van activiteiten, zijn meestal gerelateerd aan de specifieke domeinkennis van de lezer [25]. Dit soort factoren worden daarom ook niet vaak aangehaald in de bestaande literatuur. Ook factoren rond het *modeldoel* worden nauwelijks tot niet aangehaald in de literatuur.

De *modelleringsnotatie* is wel een veelbesproken factor in de literatuur. Zo werd er in een vergelijkend onderzoek, tussen verschillende modelleringsnotaties, bij 103 studenten die geen voorkennis hadden van grafische notaties ontdekt dat bij complexere modellen de voorkeur wordt gegeven aan BPMN [13]. In andere vergelijkende onderzoeken werd bovendien ontdekt dat BPMN een notatie is die meestal makkelijker begripbaar is voor alle stakeholders binnen een bedrijf [1, 11]. Dit is een belangrijk kenmerk van BPMN, aangezien het gebruik van

procesmodellen een groot voordeel kan opleveren bij de communicatie in bedrijven, als deze modellen begrijpbaar zijn voor heel de onderneming [9]. Ook de vergelijking van de begrijpbaarheid van BPMN-modellen t.o.v. tekstuele representaties is een veelbesproken onderwerp in de literatuur. Een onderzoek hierover bij studenten en beoefenaars van BPMN toonde aan dat ervaren deelnemers de voorkeur geven aan BPMN-modellen, terwijl er voor onervaren deelnemers geen verschil is tussen beide representaties [5]. Een ander onderzoek, specifiek uitgevoerd op bedrijfsmanagers, toonde aan dat bij deze doelgroep een betere begrijpbaarheid geconstateerd kan worden wanneer er gebruik wordt gemaakt van tekstuele representaties [44]. In het algemeen is de meest effectieve manier echter het toevoegen van beide soorten representaties [6, 28].

Ook de *visuele lay-out* heeft een grote impact op de begrijpbaarheid van procesmodellen. Zo werd er recent aangetoond dat BPMN-modellen met een horizontale lay-out makkelijker begrijpbaar zijn dan BPMN-modellen met een verticale lay-out [16]. Bovendien wordt in de theorie vaak het gebruik van modellen die gemodelleerd zijn van links naar rechts (left-to-right) aangeraden, aangezien dit overeenkomt met de gebruikelijke leesrichting [8]. Echter gaf een onderzoek, dat werd uitgevoerd in 2015, minder eenduidige resultaten hierover weer [8]. Hier werd namelijk geconcludeerd dat lezers zich ook goed kunnen aanpassen aan ongebruikelijke leesrichtingen [8]. Een andere factor van de visuele lay-out van procesmodellen is het kruisen van de verschillende flows. Volgens onderzoek heeft dit een beduidende negatieve impact op de begrijpbaarheid van procesmodellen [31]. In het algemeen is het belangrijk om te zorgen voor een goede en duidelijke lay-out om procesmodellen begrijpbaar te maken [6].

Onderstaande tabel (*Tabel 1*) vat enkele van deze factoren samen. De eerste kolom bevat de vijf meest besproken hoofdgroepen, de tweede kolom omvat enkele voorbeelden van factoren en de derde kolom lijst ten slotte enkele referenties op waarin de desbetreffende factoren bestudeerd worden.

	Factoren	Referenties
Model gerelateerde factoren	Cross-connectiviteit	[36]
	Aantal verbindingen	[36]
	Aantal AND-joins	[36]
	Modelgrootte	[4, 13, 23, 24, 36, 42]
Persoonlijke factoren	Praktische ervaring	[25, 36]
	Theoretische kennis	[25, 36]
Tekstuele inhoud	Activiteit-namen	[25]
Modeleringsnotatie	Modeleringsnotaties	[1, 11, 13]
	Tekstuele representatie	[5, 6, 28, 44]
Visuele lay-out	Horizontale/Verticale lay-out	[8, 16]
	Left-to-right/Right-to-left lay-out	[8]
	Kruisingen van flows	[31]

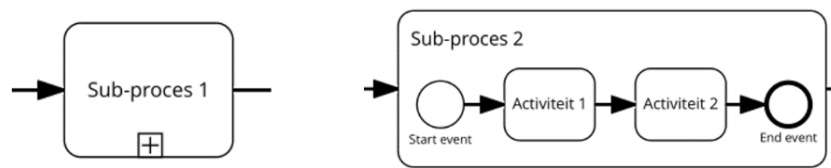
Tabel 1. Factoren van de begrijpbaarheid

2.3 Modulariteit

Zoals hierboven al kort werd toegelicht heeft de toenemende grootte van procesmodellen een belangrijke negatieve impact op de begrijpbaarheid [4, 13, 23, 24, 36, 42]. Vele onderzoeken en praktijkboeken raden een maximum aan van 7 activiteiten in een weergave van een BPMN-model [27, 42]. Bovendien blijkt uit onderzoek dat modellen met meer dan 50 activiteiten 50% kans hebben op fouten [20]. Op basis van deze inzichten wordt er vaak uitgegaan van een bovengrens van 50 activiteiten per weergave [42]. Ook het scrollen doorheen een BPMN-model heeft een negatieve impact op de begrijpbaarheid en moet daarom vermeden worden [16]. Om aan deze adviezen te voldoen en de grote van een procesweergave te verkleinen, wordt er vaak gebruik gemaakt van de *modulariteit* van procesmodellen om grote modellen op te splitsten in meerdere kleinere modellen [22, 33, 33]. Het gebruik van *modularisatie*, bv. door middel van sub-processen, wordt in het algemeen gezien als een methode om de begrijpbaarheid van procesmodellen te vergroten [40, 41].

2.3.1 Modulariteit binnen BPMN

In BPMN, de modeleringsnotatie waarop deze paper zich focust, wordt *modulariteit* hoofdzakelijk toegepast door het gebruik van sub-processen [33]. In grote lijnen bestaan er in BPMN twee soorten sub-processen. Het *collapsed sub-proces*, waarin het totaal aantal activiteiten in de hoofdweergave van het procesmodel wordt verkleind door de interne stappen van het sub-proces te verbergen [7]. Dit soort sub-proces wordt aangegeven door het sub-proces, weergegeven in de hoofdweergave, te voorzien van een vierkantje met een plus-teken erin [7]. Het sub-proces kan dan worden geraadpleegd, via het plusteken, in een ander venster [7]. En het *expanded sub-proces*, waarbij de interne stappen gewoon getoond worden in de hoofdweergave van het procesmodel [7]. Dit sub-proces wordt weergegeven door een afgeronde vierhoek met een volle lijn rond de interne stappen van het sub-proces te zetten [7]. Onderstaande afbeelding (*Figuur 1*) geeft een voorbeeld weer van een *collapsed sub-proces* (links) en een *expanded sub-proces* (rechts).



Figuur 1. Soorten sub-processen

2.3.2 Effect van sub-processen

Het toepassen van *modulariteit* door middel van sub-processen heeft twee tegenstrijdige effecten, die een grote diversiteit veroorzaken in de uitkomsten van de

verschillende reeds uitgevoerde onderzoeken [40, 41]. Langs een kant wordt, door het verbergen van minder belangrijke informatie, de nodige *mentale inspanning* voor het begrijpen van het BPMN-model kleiner, ook wel het *information hiding effect* genoemd [33, 40, 41]. Langs de andere kant wordt de nodige *mentale inspanning* groter aangezien de lezer zijn aandacht moet verdelen over meerdere fragmenten en moet bladeren tussen de gefragmenteerde stukken, ook wel het *split-attention effect* of de *browsing cost* genoemd [33, 40, 41].

2.3.3 Resultaten uit voorgaande onderzoeken

In 2016 voerden Turetken et al. een onderzoek uit naar de begrijpbaarheid van BPMN-modellen op 60 mensen uit het bedrijfsleven met een bepaalde kennis van procesmodellen, maar slechts een beperkte kennis van BPMN specifiek [41]. In dit onderzoek werd gekeken naar de begrijpbaarheid van drie verschillende vormen van *modulariteit*: fully flattened modellen, fully flattened modellen waarin bepaalde activiteiten gegroepeerd zijn m.b.v. groups en modellen met *collapsed sub-processen* [41]. Turetken et al. konden concluderen dat voor mensen uit het bedrijfsleven de begrijpbaarheid het hoogste was bij fully flattened modellen zonder groups [41].

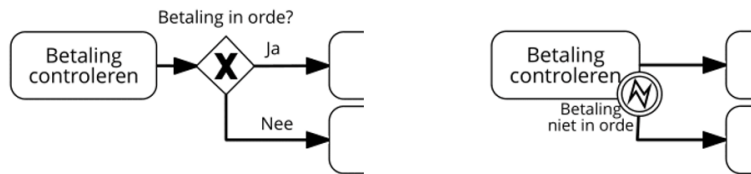
In 2020 gingen Turetken et al. dieper in op de bovenstaande gevonden resultaten door hetzelfde onderzoek uit te voeren op een tweede groep participanten van 140 afgestudeerde studenten die deelnamen aan een cursus rond bedrijfsprocesbeheer [40]. Na het samenvoegen van beide onderzoeken kon geconcludeerd worden dat ook voor de pas afgestudeerde studenten, zonder praktische kennis van procesmodellen, de *effectiviteit* daalt door het gebruik van sub-processen zonder dat het bijdraagt aan de *efficiëntie* [40]. De *effectiviteit* wijst in dit geval op de mate waarin de participanten de vragen over het procesmodel correct konden beantwoorden [40]. De *efficiëntie* wijst op de tijd die de participanten nodig hadden om de vragen juist te beantwoorden [40]. Turetken et al. merkte hierbij ook enkele beperkingen van hun onderzoeken op [40]. Zo werd onder andere het gebruik van *expanded sub-processen* niet getest en werden de onderzoeken enkel uitgevoerd op procesmodellen met meerdere pools [40].

Het empirische onderzoek van Reijers et al. over de begrijpbaarheid van Workflow Nets op 28 ervaren procesmodelleurs, toonde aan dat subprocessen, door het gevolg van het *information hiding* effect, een positief effect hebben op de begrijpbaarheid van procesmodellen [33–35]. Dit positieve effect kwam echter wel enkel tot uiting wanneer er uitgebreid gebruik werd gemaakt van sub-processen [33–35].

2.4 Conditities

Het voorkomen van *condities* of beslissingen in je proces kan in sommige omstandigheden op twee manieren gemodelleerd worden. Langs een kant kan er gewerkt worden met een *exclusive gateway*, ook wel een *XOR-gateway* genoemd. *XOR-gateways* worden gebruikt wanneer er twee of meer alternatieve wegen zijn in

een procesmodel en er een bepaalde beslissing genomen moet worden of een bepaalde *conditie* gecontroleerd moet worden [32, 43]. Een *XOR-gateway* kan werken op twee verschillende manieren. Wanneer een *XOR-gateway* één inkomende sequence flow en meerdere uitgaande sequence flows heeft, zal het de sequence flow splitsen op basis van een bepaalde beslissing of *conditie* [37, 43]. Wanneer een *XOR-gateway* meerdere inkomende sequence flows, maar slechts één uitgaande sequence flow heeft, zal het de inkomende sequence flows samenvoegen tot één uitgaande sequence flow [37, 43]. De *XOR-gateway* wordt aangeduid door een lege diamant vorm of een diamant vorm met een 'X' erin [37, 43]. Langs de andere kant kan er ook gebruik gemaakt worden van een *intermediate error event*. Dit is een event dat gekoppeld is aan de grens van een activiteit [43]. Wanneer er een error (of een bepaalde *conditie*/beslissing) voorkomt, onderbreekt dit event de activiteit waaraan het bevestigd is [43]. Het *intermediate error event* wordt aangeduid met een bliksemschicht omcirkelt door twee cirkels met een volle lijn [43]. Onderstaande afbeelding (*Figuur 2*) illustreert een situatie waarin de *XOR-gateway* vervangen kan worden door een *intermediate error event*. Aan de linker kant wordt een *XOR-gateway* weergegeven, aan de rechterkant wordt deze *XOR-gateway* vervangen door een *intermediate error event*.



Figuur 2. Weergegeven van condities

2.4.1 Resultaten uit voorgaande onderzoeken

In 2013 voerde Recker een empirisch onderzoek uit waarin de begrijpbaarheid van modellen met gateways vergeleken werden met de begrijpbaarheid van modellen zonder gateways [32]. Dit onderzoek toonde aan dat gateways helpen bij de interpretatie van procesmodellen en een positief effect hebben op de begrijpbaarheid [32].

Bovendien toonde een ander onderzoek van Sánchez-González et al. in 2011 aan dat, wanneer er rekening gehouden wordt met de begrijpbaarheid van procesmodellen, een procesmodel niet meer dan 31 gateways zou mogen bevatten [39]. Daarbij zouden 21 van deze 31 gateways *XOR-gateways* mogen zijn [39].

2.5 Activiteit-namen

Een procesmodel bevat onder andere *syntactische* en *semantische* informatie. De *syntactische* informatie verwijst naar de structuur en het gedrag van een

proces [12, 25]. De *semantische* informatie verwijst naar de tekstuele inhoud weergegeven in het procesmodel, die nodig is om duidelijk te maken aan de lezer waar het model over gaat. [12]. Onder andere de *activiteit-namen* maken onderdeel uit van deze *semantische* informatie. Zo bevatten activiteiten die gelabeld zijn met hoofd letters enkel *syntactische* informatie, terwijl activiteiten met echte namen zowel *syntactische* als *semantische* informatie bevatten [25]. Het weergeven van deze *semantische* informatie, door het gebruik van labels of echte *activiteit-namen*, heeft impact op de begrijpbaarheid van procesmodellen.

2.5.1 Effect van activiteit-namen

Het gebruik van *activiteit-namen* in procesmodellen heeft twee tegenstrijdige effecten. Langs een kant is een model makkelijker begrijpbaar wanneer er gewerkt wordt met echte *activiteit-namen*, aangezien het logische verloop van het proces afgeleid kan worden uit de opeenvolgende activiteiten [25]. Kennis van het domein en de terminologie van een proces kan bovendien ook helpen bij het begrijpen van procesmodellen [40]. Langs de andere kant beschrijft de *cognitieve belasting theorie (CLT)* dat er tijdens het verwerken van nieuwe informatie rekening moet worden gehouden met een gelimiteerd werkgeheugen en een praktisch ongelimiteerd lange-termijn geheugen [25, 29]. Wanneer mensen nieuwe informatie binnenkrijgen, wordt deze informatie altijd eerst verwerkt door het werkgeheugen, vooraleer het opgeslagen wordt in het lange-termijn geheugen [29]. Het verwerken van informatie weergegeven in een procesmodel gebeurt dus eerst door het werkgeheugen [25]. Echte *activiteit-namen* zorgen hierbij voor een hogere cognitieve belasting aangezien er meer werkgeheugen nodig is om de tekstuele informatie van de *activiteit-namen* te verwerken [25]. Dit kan op zijn beurt de persoon weerhouden om juiste conclusies te trekken over de *syntactische* aspecten van het procesmodel [25].

2.5.2 Resultaten uit voorgaande onderzoeken

In 2012 voerden Mendling et al. een onderzoek uit waarin een enquête werd opgesteld rond meerdere procesmodellen [25]. Sommige van deze procesmodellen bevatten activiteiten met echte namen, sommige procesmodellen bevatten activiteiten met abstracte namen (door ze te labelen met hoofdletters) [25]. Hieruit bleek dat extra *semantische* informatie in de vorm van tekstuele *activiteit-namen*, een negatief effect heeft op de begrijpbaarheid van de *syntactische* inhoud van een proces model [25]. Zowel *understandability task effectiveness* als *understandability task efficiency* waren hoger wanneer er gewerkt werd met abstracte labels [25].

In een ander onderzoek van Mendling et al. werd er door middel van verschillende theorieën aangetoond dat de begrijpbaarheid van procesmodellen verbeterd kan worden door de *activiteit-namen* te complementeren met iconische grafische representaties [21].

Bovendien raden verschillende richtlijnen, die opgesteld werden in de literatuur, de werkwoord-naamwoord stijl aan voor *activiteit-namen* [22, 38]. Volgens deze stijl zouden *activiteit-namen* moeten beginnen met een werkwoord in de ge-

biedende wijs dat verwijst naar de activiteit die uitgevoerd wordt, gevolgd door het bedrijfsobject dat bij de activiteit betrokken is [14]. Dit werd ook onderzocht in een onderzoek uitgevoerd door Mendling et al. in 2010, waarin werd aangetoond dat *activiteit-namen* in werkwoord-naamwoord stijl een positievere impact hebben op de begrijpbaarheid t.o.v. andere stijlen [19].

3 Probleemstelling

De begrijpbaarheid van procesmodellen kan afhangen van veel verschillende factoren: de modeleringsnotatie, de procesmodel grote, de kennis van lezer, etc.. De tegenstrijdige resultaten van voorgaande onderzoeken tonen aan hoe complex de relatie is tussen deze verschillende factoren [6]. Door deze complexe relatie en de stijgende populariteit van procesmodellen is er nood aan extra onderzoek naar de impact van deze factoren in verschillende omstandigheden [6]. Het doel van deze paper is om extra duidelijkheid te bieden over de impact van enkele van deze factoren op de begrijpbaarheid van procesmodellen.

De eerste factor die bestudeerd wordt is de *modulariteit* van BPMN-modellen. Een literatuurstudie uitgevoerd in 2018 toont onder andere aan dat er een tekort is aan onderzoeken rond deze factor [6]. Om in te spelen op de beperkingen van de reeds uitgevoerde onderzoeken van Turetken et al., die werden besproken in sectie 2.3, wordt er in deze paper specifiek gekeken naar het verschil tussen het gebruik van *expanded sub-processen* en *collapsed sub-processen* bij procesmodellen met slechts 1 pool [40,41]. Bovendien wordt in de reeds bestaande onderzoeken de domeinkennis van de participanten niet buiten beschouwing gehouden [40]. In deze paper wordt hier, door te werken met *abstracte activiteit-namen*, wel rekening mee gehouden [32]. Een tweede factor van de begrijpbaarheid van procesmodellen die bekeken wordt is het gebruik van *XOR-gateways* t.o.v. het gebruik van *intermediate error events*. Ook hierbij wordt er gewerkt met *abstracte activiteit-namen* om de domeinkennis van participanten buiten beschouwing te laten. De derde en laatste factor die wordt opgenomen in deze paper is een factor rond de tekstuele inhoud van procesmodellen en omvat de *activiteit-namen*. Er wordt dieper ingegaan op de impact van *activiteit-namen* door te kijken naar BPMN-modellen met *abstracte activiteit-namen*, tekstueel *logische activiteit-namen* en tekstueel *onlogische activiteit-namen*.

4 Methodologie

In deze paper wordt onderzocht hoe het gebruik van *modulariteit*, het gebruik van *XOR-gateways* en *intermediate error events* en het gebruik van *activiteit-namen* de begrijpbaarheid van procesmodellen in BPMN beïnvloed. Om de impact van deze factoren te bestuderen werd er gekozen voor een empirisch onderzoek waarin wordt gewerkt met primaire data die wordt verzameld via een enquête. Deze enquête bevat vragen over enkele statische procesmodellen die werden gemodelleerd in BPMN 2.0 via Signavio (www.signavio.com). Vanuit de verzamelde data wordt de begrijpbaarheid van de gebruikte procesmodellen geoperationaliseerd

m.b.v. twee indicatoren: de *understandability task effectiveness* en de *mentale inspanning*.

Binnen deze sectie wordt de opzet van het voorgestelde onderzoek verder uiteengezet. Allereerst worden de onderzoeksvragen opgesteld en worden de onafhankelijke variabelen aangehaald. Bovendien worden ook de procesmodellen, die gebruikt worden om de onafhankelijke variabelen te operationaliseren, kort beschreven. Hierna worden ook de afhankelijke variabelen kort aangehaald. In sectie 4.6 wordt vervolgens uitgelegd hoe het onderzoek onderverdeeld werd in verschillende onderzoeksblokken. Sectie 4.7 beschrijft hoe de vragen van de enquête werden samengesteld. Vervolgens worden in sectie 4.8 de participanten van het onderzoek kort beschreven. Ten slotte wordt in sectie 4.9 uitgelegd hoe er verder met de verzamelde data uit de enquête gewerkt werd.

4.1 Onderzoeksvragen

Gebaseerd op bovenstaande informatie werden en drie onderzoeksvragen opgesteld. In de eerste onderzoeksvraag wordt er gekeken naar de *modulariteit* van BPMN-modellen. De tweede onderzoeksvraag bestudeert het modelleren van *condities* a.d.h.v. *XOR-gateways* en *intermediate error events*. De derde en laatste onderzoeksvraag gaat over de *activiteit-namen* binnen een procesmodel.

Onderzoeksvraag 1 Is het gebruik van *expanded* of *collapsed sub-processen* beter voor de begrijpbaarheid van BPMN-modellen?

Onderzoeksvraag 2 Is het gebruik van *XOR-gateways* of *intermediate error events* beter voor de begrijpbaarheid van BPMN-modellen?

Onderzoeksvraag 3 Wat is de impact van *activiteit-namen* op de begrijpbaarheid van BPMN-modellen wanneer er gekeken wordt naar *abstracte activiteit-namen*, *logische activiteit-namen* en *onlogische activiteit-namen*?

4.2 De onafhankelijke variabele: Modulariteit

Om een antwoord te geven op de eerste onderzoeksvraag werd de factor *modulariteit* geoperationaliseerd als een onafhankelijke variabele door twee verschillende versies te maken van een procesmodel. Eerst werd er een procesmodel ontwikkeld met meerdere *expanded sub-processen* (*procesmodel 1A*). Vervolgens werd er een tweede versie van dit procesmodel gemaakt waarin alle *expanded sub-processen* vervangen werden door *collapsed sub-processen* (*procesmodel 1B*). Beide procesmodellen kunnen teruggevonden worden in de bijlagen (*Bijlage A en B*). Om de impact van de domeinkennis van de participanten en het gebruik van *activiteit-namen* buiten beschouwing te houden, werd er in beide procesmodellen gewerkt met *abstracte activiteit-namen* m.b.v. hoofdletters [32]. Deze hoofdletters werden op volgorde van boven naar onder in verschillende kolommen over het procesmodel verdeeld. Op die manier is het makkelijker voor de participanten om de

hoofdletters terug te vinden, maar zat er geen logica van het proces in de volgorde van de hoofdletters. Verder werd er gekozen om te werken met procesmodellen die ongeveer 20 activiteiten bevatten, aangezien de grootte van procesmodellen in de praktijk ook vaak rond dit gemiddelde gesitueerd is [14, 30].

4.3 De onafhankelijke variabelen: Conditie

De tweede onderzoeksvraag werd beantwoord op een gelijkaardige manier als de eerste. Opnieuw werden er twee versies gemaakt van een bepaald procesmodel. In de eerste versie wordt gebruik gemaakt van meerdere *XOR-gateways* (*procesmodel 2A*). In de tweede versie worden de *condities* weergegeven door deze *XOR-gateways* te vervangen door *intermediate error events* (*procesmodel 2B*). Ook deze procesmodellen kunnen teruggevonden worden in de bijlagen (*Bijlage C en D*). Bovendien werd ook hier gekozen voor procesmodellen met ongeveer 20 activiteiten, waarbij gebruik wordt gemaakt van *abstracte activiteit-namen* m.b.v. hoofdletters die opnieuw op volgorde van boven naar onder in verschillende kolommen over het procesmodel werden verdeeld.

4.4 De onafhankelijke variabele: Activiteit-namen

De derde en laatste onderzoeksvraag gaat over de impact van *activiteit-namen* op de begrijpbaarheid van BPMN-modellen. Om de factor *activiteit-namen* te operationaliseren als een onafhankelijke variabelen werd er gekozen om 3 verschillende versies te maken van een bepaald procesmodel, met telkens andere soorten *activiteit-namen*. De eerste versie is een procesmodel met *abstracte activiteit-namen* m.b.v. hoofdletters (*procesmodel 3*). Ook bij dit procesmodel werd er rekening gehouden met een modelgrootte van ongeveer 20 activiteiten, zoals de meeste procesmodellen die in de praktijk voorkomen [14, 30]. Opnieuw werden de hoofdletters verdeeld over het procesmodel op volgorde van boven naar onder in verschillende kolommen. In de tweede versie van het procesmodel werden simpelweg alle *abstracte activiteit-namen* vervangen door *logische activiteit-namen* (*procesmodel 4A*). Om ervoor te zorgen dat de verschillen tussen de participanten op vlak van domein kennis minimaal blijven, werd er gekozen voor een procesmodel dat algemeen gekend is door de meeste mensen, namelijk een procesmodel dat hotel-roomservice beschrijft. Dit procesmodel werd gebaseerd op een procesmodel dat werd gevonden in de literatuur [10]. Additioneel werd er vervolgens ook nog een derde versie aangemaakt van het procesmodel. In deze versie werden enkele van de activiteiten van *procesmodel 4A* van plaats verwisseld, waardoor er een procesmodel ontstond met *onlogische activiteit-namen* (*procesmodel 4B*). Ook deze procesmodellen kunnen teruggevonden worden in de bijlage (*Bijlage E, F en G*).

4.5 De afhankelijke variabelen

Zoals werd toegelicht in sectie 2 kunnen er meerdere soorten indicatoren gebruikt worden om de begrijpbaarheid van BPMN-modellen te operationaliseren.

In dit onderzoek werd er gekozen om dit te doen via de *understandability task effectiveness* en de *mentale inspanning*.

De *understandability task effectiveness* is een objectieve indicator die in dit onderzoek wordt gemeten door een aantal vragen te stellen over de voorgelegde procesmodellen. Vervolgens wordt het aantal correct beantwoorde vragen gedeeld door het totaal aantal vragen om de *understandability task effectiveness* score te berekenen.

De *mentale inspanning* is een subjectieve indicator die in dit onderzoek wordt gemeten door na elke begrijpbaarheidsvraag over het procesmodel te vragen naar de waargenomen moeilijkheidsgraad. De participanten geven hier hun mening over de moeilijkheid van de gestelde begrijpbaarheidsvraag.

4.6 De onderzoeksblokken

Om bovenstaande onafhankelijke variabelen te bestuderen a.d.h.v. de afhankelijke variabele wordt de enquête opgedeeld in vijf verschillende blokken, van blok 0 tot en met blok 4.

BLOK 0 bevat enkele inleidende vragen over de demografische gegevens van de participanten.

BLOK 1 gaat over de *modulariteit* van procesmodellen (*onderzoeksvraag 1*). Hiervoor worden de participanten random over 2 experimentele groepen verdeeld. Een groep krijgt vragen over het procesmodel met de *expanded sub-processen* (*procesmodel 1A*), de andere groep krijgt vragen over het procesmodel met de *collapsed sub-processen* (*procesmodel 1B*).

BLOK 2 gaat vervolgens over het modelleren van *condities* in procesmodellen (*onderzoeksvraag 2*). Ook hier worden de participanten opnieuw random over 2 groepen verdeeld waarbij een groep vragen krijgt over het procesmodel met de *XOR-gateways* (*procesmodel 2A*) en de andere groep vragen krijgt over het procesmodel met de *intermediate error events* (*procesmodel 2B*).

BLOK 3 en 4 gaan ten slotte over de activiteit-namen in procesmodellen (*onderzoeksvraag 3*). Voor deze onderzoeksvraag werden 3 verschillende modellen gemaakt. In BLOK 3 krijgen alle participanten eerst vragen over het procesmodel met de *abstracte activiteit-namen* (*procesmodel 3*). Vervolgens, in BLOK 4 worden de participanten weer random verdeeld over 2 groepen. Een groep krijgt vragen over *procesmodel 4A*, om te kijken wat de impact is van het toevoegen van *logische activiteit-namen*. De andere groep krijgt vragen over *procesmodel 4B*, om te kijken wat de impact is van het toevoegen van *onlogische activiteit-namen*.

4.7 De enquête

De enquête werd gemaakt via de software Qualtrics. Om de enquête toegankelijk te maken voor participanten van verschillende afkomst, werd er gekozen om de vragen in het engels weer te geven. In totaal waren er 3 soorten vragen: de inleidende vragen, de begrijpbaarheidsvragen en de moeilijkheidsvragen.

4.7.1 Inleidende vragen

In het eerste deel van de enquête (BLOK 0) wordt er gevraagd naar de huidige situatie van de participant (student/werk), het opleidingsniveau, de kennis van het modelleren van processen (geen kennis/theoretische kennis/praktische kennis) en de kennis van BPMN als modeleringsnotatie specifiek (op een likert-schaal van "I never heard of it" (1) t.e.m. "I'm an expert" (5)). De kennis van participanten is een belangrijk element aangezien onderzoeken in het verleden al hebben aangetoond dat persoonlijke kenmerken, zoals praktisch ervaring met of theoretische kennis van proces modelleren, vaak zelfs meer impact op de begripbaarheid hebben dan de model gerelateerde kenmerken [36]. Bovendien kregen de participanten aan het begin van de enquête het advies om de enquête in te vullen via een pc of laptop. De kleine weergave van een gsm en de mogelijke verplichting tot scrollen doorheen het model (*split-attention effect*) kan de resultaten van de enquête namelijk beïnvloeden [16, 33, 40, 41]. Om meer zicht te hebben op het uiteindelijk gebruikte toestel wordt er voor de zekerheid als laatste gevraagd op welk toestel de participant de enquête invult. Deze vragen en de mogelijke antwoord opties kunnen ook in de bijlage teruggevonden worden (*Bijlage H*).

4.7.2 Begrijpbaarheidsvragen

De begripbaarheidsvragen dienen om de *understandability task effectiveness* score te berekenen en te evalueren hoe goed de participanten de gegeven procesmodellen begrijpen. Per procesmodel werden er 6 begripbaarheidsvragen gesteld met telkens een aantal antwoord opties waaruit de participant kon kiezen (meerkeuzevragen). Bovendien is er ook altijd een extra antwoordoptie "I don't know" ter beschikking. In BLOK 1 en 2 kregen de twee onderzoeksgroepen exact dezelfde vragen, maar een andere versie van het procesmodel. In BLOK 3 kreeg de gehele onderzoeksgroep hetzelfde procesmodel met dezelfde vragen. In BLOK 4 kregen de twee onderzoeksgroepen andere versies van het procesmodel met *syntactisch* dezelfde vragen als in BLOK 3. De *semantische* betekenis van de vragen is hier echter anders, aangezien de *activiteit-namen* ook *semantisch* anders waren. Bovendien werd in de enquête de volgorde van de vragen veranderd ten opzichte van de vragen in BLOK 3, zodat dit minder makkelijk met elkaar te vergelijken was. Participanten hadden hier bovenop niet de optie om terug te gaan naar het vorige procesmodel en konden dus de antwoorden niet zomaar overnemen door terug te kijken. Als extra controle kregen alle participant bovendien ook 2 test-vragen tussen de begripbaarheidsvragen. In deze test-vragen werd aan de participanten gevraagd om een bepaalde antwoordoptie aan te duiden. Op deze manier kon er achteraf gecontroleerd worden of er participanten waren die de enquête random ingevuld hebben, zonder de vragen aandacht te bekijken. Ook de begripbaarheidsvragen kunnen geraadpleegd worden in de bijlage (*Bijlage I, J, K, L en M*).

4.7.3 De moeilijkheidsvragen

Na elke begripbaarheidsvraag werd er een vraag gesteld over de moeilijkheid van de desbetreffende vraag. De participanten moeten hier telkens aangeven hoe moeilijk ze de gegeven begripbaarheidsvraag vonden aan de hand van een likert-schaal. Deze likert-schaal bevat telkens 5 antwoord opties die lopen van makkelijk (1) tot en met moeilijk (5).

4.7.4 Test

Voordat de enquête effectief werd uitgestuurd naar mogelijke participanten werd de enquête eerst getest op een student. Door deze test kon er een tijd van 25 minuten geplakt worden op het invullen van de enquête. Ook werden naar aanleiding van deze test nog enkele kleine foutjes uit de enquête gehaald. Bovendien werd de volgorde waarin de verschillende onderzoeksblokken aan bod kwamen aangepast. Er werd beslist om volgende volgorde aan te nemen: eerst BLOK 2, vervolgens BLOK 3, BLOK 1 en als laatste BLOK 4. Op deze manier komen de processen van BLOK 4 met de *logische* en *onlogische activiteit-namen* niet rechtstreeks na het proces van BLOK 3 met de *abstracte activiteit-namen*, waardoor het minder evident is om hier antwoorden van over te nemen.

4.8 De participanten

De enquête werd verstuurd naar alle studenten via de mailinglijst van de UHasselt. Bovendien werd de link via andere media ook verdeeld onder verschillende groepen buiten de universiteit, om op deze manier ook mensen te betrekken zonder diploma en/of kennis van proces modelleren of BPMN. Om mensen aan te moedigen om de enquête in te vullen ontvingen de participanten op het einde van de enquête de uitslag van hun prestatie. Dit gebeurde in de vorm van een score op 6 voor elk van de 4 procesmodellen. Uiteindelijk waren er 132 participanten die de enquête hebben gestart, waarvan 70 participanten de enquête volledig hebben afgerond. Om te verzekeren dat deze participanten de enquête aandachtig hadden ingevuld, werden de test-vragen gecontroleerd. 4 participanten gaven één of meerdere fout(en) antwoorden op de test-vragen en werden daarom uit de resultaten weggefilterd. Zo bleven er uiteindelijk 66 definitieve participanten over.

4.9 De analyse

Vanuit de verzamelde data wordt de begripbaarheid van de gebruikte procesmodellen geoperationaliseerd m.b.v. twee indicatoren: de *understandability task effectiveness* en de *mentale inspanning*. Aangezien er per procesmodel 6 begripbaarheidsvragen werden gesteld, wordt de *understandability task effectiveness* telkens weergegeven als een score op 6. Hoe hoger de *understandability task effectiveness*, hoe beter de begripbaarheid. Vervolgens werd er na elke begripbaarheidsvraag gevraagd naar de *mentale inspanning* van de participanten m.b.v. een likert-schaal van makkelijk (1) tot en met moeilijk (5). De *mentale*

inspanning wordt daarom berekend als een score op 30 per procesmodel. Hier is de redenering juist omgekeerd: hoe lager de aangegeven *mentale inspanning* van de participanten, hoe begrijpbaarder het procesmodel.

Vooraleer de onderzoeksvragen beantwoord kunnen worden, is het belangrijk om de data te analyseren. In dit onderzoek gebeurt dit a.d.h.v. de statistische software R. Allereerst moet er worden gecontroleerd of de metingen van de twee afhankelijke variabelen over de verschillende onafhankelijke variabelen normaal verdeeld zijn [3]. Aangezien de steekproefgrootte per procesmodel meestal rond de 30 participanten ligt (behalve bij *procesmodel 3*), werd er gekozen om de *Shapiro-Wilk test* te gebruiken. Deze test is namelijk meer geschikt voor de analyse van data met een kleinere steekproefgrootte [26]. Zoals in de meeste experimentele onderzoeken wordt er verder gewerkt met een significantieniveau van 0.05.

Vervolgens wordt er een bijpassende statistische test gekozen om te controleren of er een significant verschil is tussen de begrijpbaarheid van de verschillende procesmodellen. Indien de data normaal verdeeld is, wordt er gewerkt met een parametrische test [3, 18, 26]. De data van *procesmodel 1A* en *procesmodel 1B* zijn onafhankelijk van elkaar. Ook de data van *procesmodel 2A* en *2B*) en de data van *procesmodel 4A* en *4B* zijn onafhankelijk van elkaar. Voor deze onafhankelijke data wordt er in dit geval gebruik gemaakt van de *Independent Samples t-test* [3, 18]. De data tussen *procesmodel 3* en *4A* of *4B* zijn afhankelijk van elkaar, aangezien de participanten ofwel *procesmodel 3* en *4A*, ofwel *procesmodel 3* en *4B* hebben getest. Voor deze afhankelijke data moet er in dit geval gewerkt worden met de *Paired samples t-test* [3, 18]. Als de data niet normaal verdeeld is, wordt er gewerkt met een niet-parametrische test die geen assumpties maakt over de verdeling van de data [3, 18, 26]. In dit geval wordt voor de onafhankelijke data de *Mann-Whitney U-test* en voor de afhankelijke data de *Wilcoxon Signed rank test* gebruikt om te controleren of er een significant verschil is tussen de verschillende versies van de procesmodellen [3, 18].

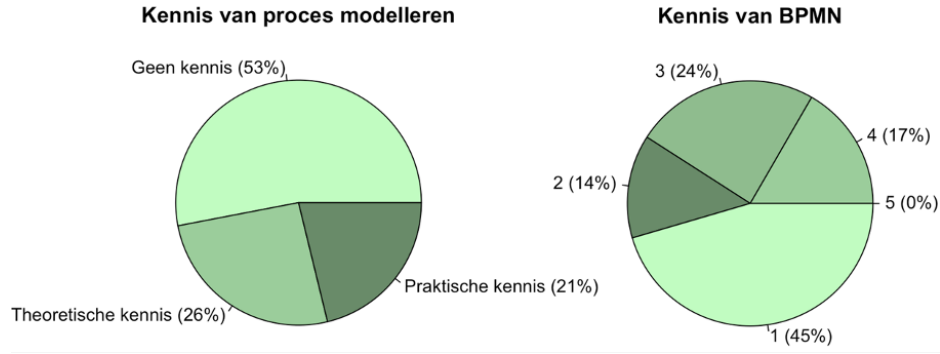
5 Resultaten en discussie

In deze sectie wordt de verzamelde data geanalyseerd en worden de resultaten besproken. In sectie 5.1 wordt de verzamelde data kort beschreven. In sectie 5.2 worden vervolgens inferentiële statistieken gebruikt om een antwoord te geven op de drie onderzoeksvragen van deze paper en worden de resultaten kort besproken. Als laatste wordt er ook dieper ingegaan op de impact van de kennis van de participanten op de begrijpbaarheid.

5.1 Beschrijvende statistieken

Van de 66 participanten gaven 44 participanten aan nog student te zijn. Bovendien zijn 21 participanten in het bezit van een masterdiploma, 29 participanten in het bezit van een bachelor diploma, 14 participanten hebben als hoogst behaalde diploma een diploma van de middelbare school en 2 participanten hebben

geen diploma. Ook de kennis van het modelleren van processen en de kennis van de modeleringsnotatie BPMN is verdeeld over de participanten. In onderstaande figuur (*Figuur 3*) wordt de kennis van de participanten grafisch weergegeven. Het grootste deel van de participanten (53%) gaf aan geen kennis te hebben van het modelleren van processen. Bovendien werd er ook gevraagd naar de kennis van BPMN a.d.h.v. een likert-schaal van "I never heard of it" (1) t.e.m. "I'm an expert" (5)). Ook hier gaf het grootste deel van de participanten (45%) aan nog nooit gehoord te hebben van de modelleringsnotatie BPMN.



Figuur 3. Kennis van participanten

Alle 66 participanten kregen vragen over 4 verschillende procesmodellen. Dit resulteerde in een totaal van 264 observaties. Bovendien werd de software waarmee de enquête werd gemaakt zo ingesteld zodat de verschillende procesmodellen in alle onderzoeksblokken zo evenredig mogelijk verdeeld zouden worden over de participanten. Onderstaande tabel (*Tabel 2*) vat de beschrijvende statistieken van de gebruikte procesmodellen kort samen. Voor elk procesmodel wordt weergegeven hoeveel participanten het model hebben getest (N) en wat het gemiddelde ($\bar{x}$) en de standaard afwijking (σ) is van de *understandability task effectiveness* (UTE) en de *mentale inspanning* (MI).

Onafhankelijke variabelen		N	UTE	UTE	MI	MI
			$\bar{x}$	σ	$\bar{x}$	σ
1A	Expanded sub-processen	31	3.84	1.24	17.96	5.59
1B	Collapsed sub-processen	35	3.14	1.66	17.14	5.25
2A	XOR-gateways	34	4.91	1.33	13.15	7.04
2B	Intermediate error events	32	4.94	1.04	12.69	4.63
3	Abstracte activiteit-namen	66	4.51	1.21	12.94	5.15
4A	Logische activiteit-namen	33	4.79	1.11	11.21	4.52
4B	Onlogische activiteit-namen	33	4.61	1.25	11.61	5.02

Tabel 2. Beschrijvende statistieken

5.2 Inferentiële statistieken

Zoals reeds werd besproken in de methodologie, is het belangrijk om eerst te controleren of de data normaal verdeeld is a.d.h.v. de *Shapiro-Wilk test*. De resultaten van deze test worden weergegeven in onderstaande tabel (*Tabel 3*). Rekening houdend met een significantieniveau van 0.05 blijken enkel de verdelingen van de afhankelijke variabelen van *procesmodel 1A* en de verdeling van de *mentale inspanning* van *procesmodel 1B* normaal verdeeld zijn ($p > 0.05$). Bijgevolg werd er enkel bij de vergelijking van de *mentale inspanning* van *procesmodel 1A* en *1B* gewerkt met een parametrische test. Alle andere analyses gebeurden a.d.h.v. niet-parametrische testen.

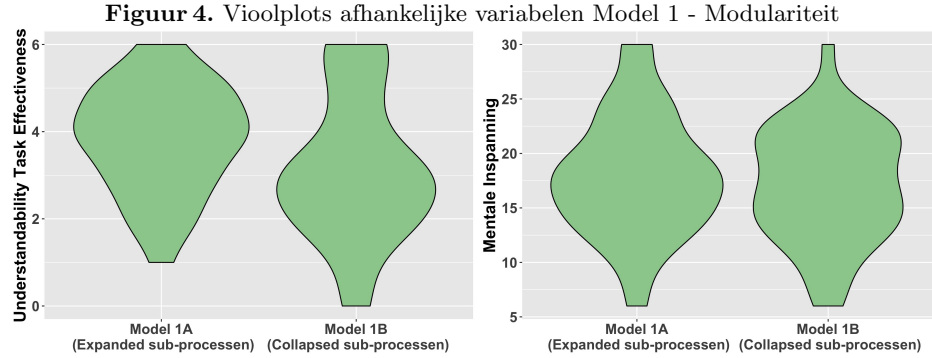
	Onafhankelijke variabelen	UTE (p)	MI (p)
1A	Expanded sub-processen	0.0521	0.6634
1B	Collapsed sub-processen	0.0128	0.4697
2A	XOR-gateways	0.0025	0.0005
2B	Intermediate error events	0.0003	0.0446
3	Abstracte activiteit-namen	0.0099	0.0155
4A	Logische activiteit-namen	0.0002	0.0036
4B	Onlogische activiteit-namen	0.0014	0.0004

Tabel 3. Shapiro-Wilk test

5.2.1 Modulariteit

Onderzoeksvraag 1 gaat over de begripbaarheid van *expanded* en *collapsed sub-processen*. *Figuur 4* geeft de vioolplots van de afhankelijke variabelen van *procesmodel 1A* (*expanded sub-processen*) en *procesmodel 1B* (*collapsed sub-processen*) weer. Als eerste werd er gekeken naar de *understandability task effectiveness*. De *Mann-Whitney U-test* resulteerde hier in een p-waarde van 0.0450. Dit toont aan dat er een significant verschil is tussen de *understandability task effectiveness* van beide procesmodellen. Ook in de vioolplots kan dit verschil duidelijk waargenomen worden. Er kan geconcludeerd worden dat een procesmodel met *expanded sub-processen* makkelijker te begrijpen is dan een procesmodel met *collapsed sub-processen*. Net zoals in het voorgaande onderzoek van Turetken et al. bevestigt dit onderzoek de voorkeur voor 'plattere' modellen [40, 41]. Het *split-attention effect* heeft kennelijk meer impact op de *understandability task effectiveness* dan het *information hiding effect*, waardoor de participanten verkiezen het volledige procesmodel in één weergave te zien a.d.h.v. *expanded sub-processen* [33, 40, 41]. Om de resultaten dieper te analyseren werd er a.d.h.v. de mediaan gekeken naar het verschil in *understandability task effectiveness* bij participanten met een andere voorkennis (kennis van proces modelleren of BPMN). Hieruit kon besloten worden dat het verschil in begripbaarheid zichtbaar is in alle kenniscategorieën. Vervolgens werd er, om te kunnen achterhalen wanneer *collapsed sub-processen*

precies voor verwarring zorgen, ook per vraag gekeken naar het percentage van participanten die een juist antwoord gaven. Op deze manier kon er gecontroleerd worden of er een specifieke vraag uit de enquête is die bij *procesmodel 1A* echt makkelijker was dan bij *procesmodel 1B*, of omgekeerd. De percentages van de vragen bij *procesmodel 1A* zijn allemaal hoger dan de percentages bij *procesmodel 1B*. De algemene begrijpbaarheid ligt bij *procesmodel 1A* dus hoger. Er zijn echter wel twee vragen waarbij dit verschil net iets groter was: vraag *Q1A-3/Q1B-3* en vraag *Q1A-6/Q1B-6* (zie *Bijlage I*). Bij *Q1A-3/Q1B-3* moesten de participanten aangeven welke activiteit niet rechtstreeks volgde op activiteit H. Bij *Q1A-6/Q1B-6* moesten de participanten aangeven hoeveel activiteiten er voorkwamen in het kortste pad naar end event 4. Dit zijn beide vragen met een grotere omvang, waarbij verschillende elementen van binnen en buiten de sub-processen samen bekeken moesten worden. Een logisch verklaring zou daarom kunnen zijn dat de aparte weergaven van het sub-proces in *procesmodel 1B*, het combineren van activiteiten die binnen en buiten de sub-processen liggen vermoeilijkt (*split-attention effect*).

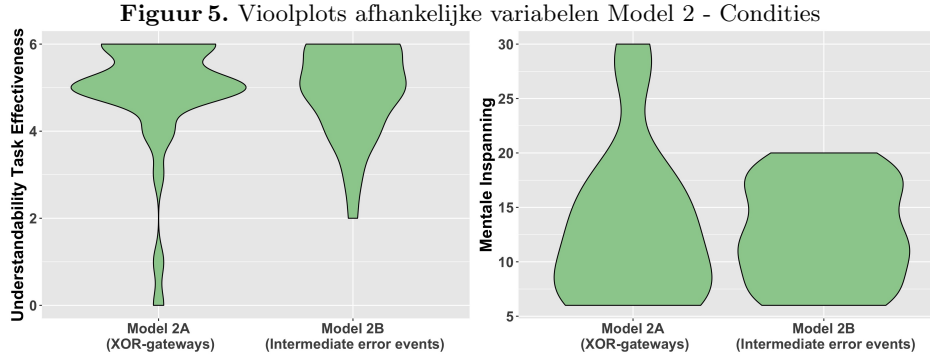


Als tweede werd er gekeken naar de *mentale inspanning*, die de participanten aangaven, bij het beantwoorden van de vragen. Ook dit kan waargenomen worden in *Figuur 4*. Het verschil tussen de vioolplots van beide procesmodellen is hier minder duidelijk. Aangezien beide verdelingen normaal verdeeld zijn, werd er gebruik gemaakt van de *Independent Samples t-test*, die resulteerde in een p-waarde van 0.5405. Dit toont aan dat er geen significant verschil is tussen de geschatte *mentale inspanning* van beide procesmodellen. Om dieper in te gaan op de resultaten werd ook hier, a.d.h.v. de mediaan, opnieuw gekeken naar het verschil in *mentale inspanning* bij participanten met andere voorkennis (kennis van proces modelleren of BPMN). Bij de participanten die geen kennis hadden van modelleren of geen kennis hadden van BPMN is de mediaan van de mentale inspanning bij *procesmodel 1A* en *1B* ongeveer gelijk. Opvallend is echter dat bij de participanten met theoretische of praktische kennis van modelleren en bij de

participanten met kennis van BPMN de mediaan van de *mentale inspanning* bij *procesmodel 1A* hoger ligt dan bij *procesmodel 1B*. Dit spreekt de bevindingen, die werden gevonden bij de *understandability task effectiveness*, tegen en kan erop wijzen dat de participanten met kennis van modelleren of kennis van BPMN de voorkeur geven aan *procesmodel 1B*, ook al maken ook zij hier meer fouten bij het beantwoorden van de begrijpbaarheidsvragen. Om dit te achterhalen werd er een nieuwe dataset gecreëerd waaruit de data van participanten zonder kennis van modelleren of kennis van BPMN werd verwijderd. De *Shapiro-Wilk test* gaf hier opnieuw aan dat de mentale inspanning normaal verdeeld is en de *Independent Samples t-test* gaf vervolgens een p-waarde terug van 0.7283. Dit toont aan dat ook hier geen significant verschil is tussen de *mentale inspanning* van beide procesmodellen en er dus geen doorgrondende conclusies gemaakt kunnen worden. Als laatste werd bovendien ook de *mentale inspanning* van elke vraag in *procesmodel 1A* en *1B* nog met elkaar vergeleken. Ook hier zijn er eveneens geen opmerkelijke bevindingen.

5.2.2 Conditities

Om een antwoord te kunnen geven op *onderzoeksvraag 2* worden in *figuur 5* de violoplots van de afhankelijke variabelen van *procesmodel 2A* (*XOR-gateways*) en *procesmodel 2B* (*intermediate error events*) weergegeven. De *Mann-Whitney U-test* gaf hier bij alle twee de afhankelijke variabelen aan dat er geen significant verschil is tussen beide procesmodellen. Hieruit kan geconcludeerd worden dat in het algemeen beide representaties even makkelijk te begrijpen zijn.

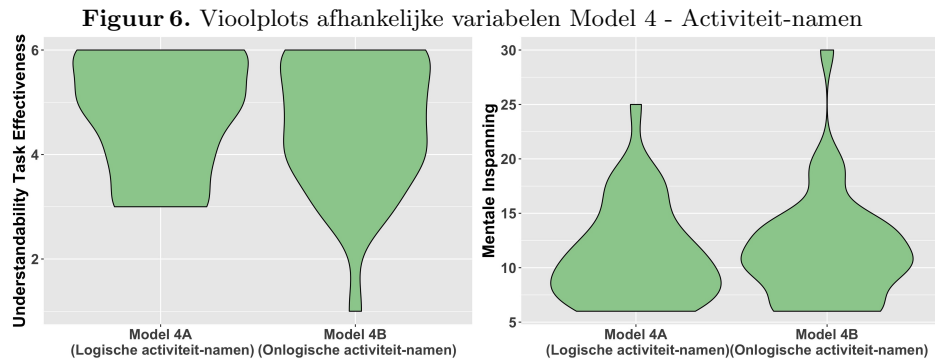


Om de resultaten verder te analyseren werd ook hier, a.d.h.v. de medianen van de afhankelijke variabelen, gekeken naar het verschil tussen participanten met andere voorkennis (kennis van proces modelleren of BPMN). Het zou namelijk kunnen voorvallen dat het niet significant verschil tussen de twee modellen veroorzaakt wordt door tegenstrijdige resultaten van de verschillende kennisgroe-

pen. De analyse toonde echter aan dat de scores van zowel de *understandability task effectiveness* als de *mentale inspanning* per kennisgroep bij beide modellen zeer gelijk lopen. Vervolgens werd er ook opnieuw gekeken naar de *understandability task effectiveness* en de *mentale inspanning* van de vragen afzonderlijk. Voor de *understandability task effectiveness* is vraag *Q2A-4 / Q2B-4* de vraag waarbij het grootste verschil tussen *procesmodel 2A* en *2B* kan worden waargenomen. Hier moesten de participanten aangeven hoeveel keer activiteit H maximaal kan voorkomen (zie *Bijlage J*). Bij *procesmodel 2B* gaf 91% van de participanten hier een goed antwoord, bij *procesmodel 2A* was dit slechts 71%. Voor de *mentale inspanning* zijn er geen opvallende verschillen tussen de vragen van *procesmodel 2A* en *2B*.

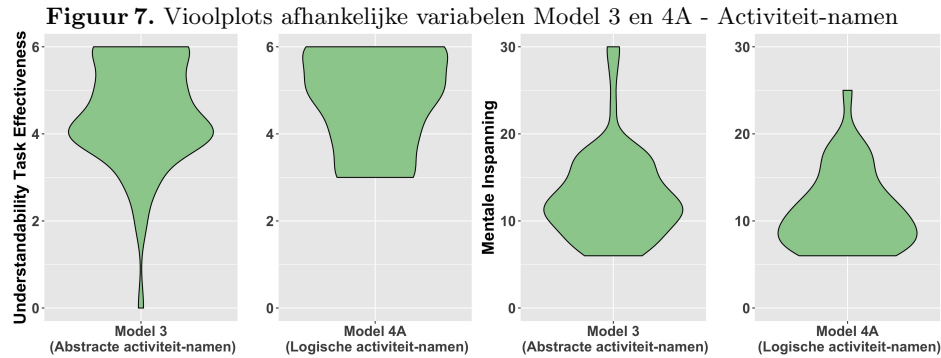
5.2.3 Activiteit-namen

Voor *onderzoeksvraag 3* werden 3 verschillende versies van een procesmodel rond hotel-roomservice getest: *procesmodel 3*, *4A* en *4B*. Allereerst werd er gekeken naar de verschillen tussen de afhankelijke variabelen van *procesmodel 4A* (*logische activiteit-namen*) en *4B* (*onlogische activiteit-namen*). *Figuur 6* geeft deze afhankelijke variabelen weer a.d.h.v. violoplots. De *Mann-Whitney U-test* van zowel de *understandability task effectiveness* als de *mentale inspanning* toonde aan dat er geen significant verschil is tussen de begrijpbaarheid van het procesmodel met *logische activiteit-namen* en het procesmodel met *onlogische activiteit-namen*.

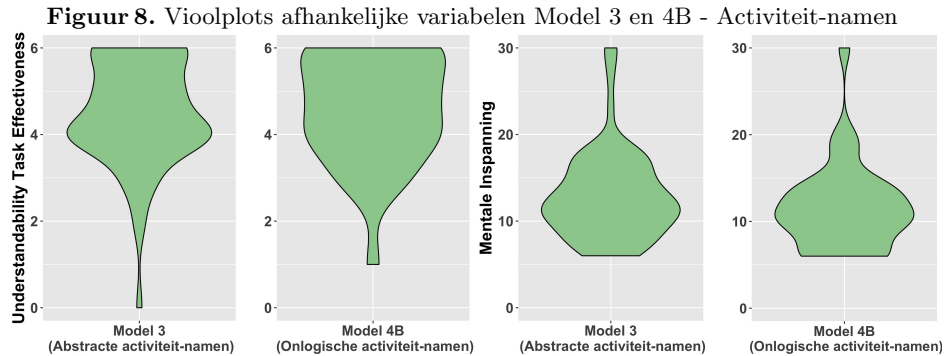


Vervolgens werd er gekeken naar de verschillen tussen de afhankelijke variabelen van *procesmodel 3* en *procesmodel 4A*. Dit werd getest met de *Wilcoxon Signed rank test*. Bij de *understandability task effectiveness* resulteerde dit in een p-waarde van 0.0498, wat aantoont dat er een significant verschil aanwezig is. Uit de violoplots in *figuur 7* kan vervolgens worden geconcludeerd dat het procesmodel met de *logische activiteit-namen* (*procesmodel 4A*) makkelijker te begrijpen is. Ook bij de *mentale inspanning* gaf de statistische test aan dat er

een significant verschil ($p=0.0003$) is tussen beide procesmodellen. Beide afhankelijke variabelen wijzen er op dat het proces model met de *logische activiteit-namen* (*procesmodel 4A*) makkelijker te begrijpen is dan het procesmodel met de *abstracte activiteit-namen* (*procesmodel 3*).



Als laatste werd er ook gekeken naar de verschillen tussen de afhankelijke variabelen van *procesmodel 3* en *4B*. Ook hier werd er gewerkt met de *Wilcoxon Signed rank test* die zowel bij de *understandability task effectiveness* als bij de *mentale inspanning* aan gaf dat het verschil tussen de afhankelijke variabelen van *procesmodel 3* en *4B* niet significant is. *Figuur 8* geeft de vioolplots van deze afhankelijke variabelen weer.



De resultaten van *procesmodel 4A* (*logische activiteit-namen*) en *4B* (*onlogische activiteit-namen*) geven aan dat er geen significant verschil is tussen de begripbaarheid van beide procesmodellen. Hieruit kan worden afgeleid dat het

gebruik van *onlogische activiteit-namen* weinig impact heeft op de begrijpbaarheid. Dit zou kunnen liggen aan twee redenen. Ofwel hielden de participanten vanaf het begin al geen rekening met de *semantische* betekenis van de *activiteit-namen*. Maar aangezien ook blijkt dat de participanten de *logische activiteit-namen* makkelijker te begrijpen vonden dan de *abstracte activiteit-namen*, is dit een beetje tegenstrijdig. Ofwel hielden de participanten geen rekening meer met de *semantische* betekenis van de *activiteit-namen* vanaf het moment dat ze door hadden dat deze onlogisch gerangschikt zijn in het procesmodel.

De resultaten van de statistische test tussen *procesmodel 3* en *4A* gaven aan dat het procesmodel met *logische activiteit-namen* (*procesmodel 4A*) beter begrijpbaar is dan het procesmodel met de *abstracte activiteit-namen* (*procesmodel 3*). Hieruit kan worden geconcludeerd dat het logische verloop van het model of de kennis van het onderwerp waarover het model gaat, meer impact heeft op de begrijpbaarheid dan de cognitieve belasting van de activiteit-namen [25, 29, 40]. Opvallend is dat dit de bevindingen van het voorgaande onderzoek van Mendling et al. tegensprekt [25]. In dit voorgaand onderzoek werd namelijk aangetoond dat de *understandability task effectiveness* en de *understandability task efficiency* alle twee hoger zijn bij het gebruik van *abstracte activiteit-namen* t.o.v. het gebruik van echte *logische activiteit-namen* [25].

Als laatste toonde ook de resultaten van *procesmodel 3* (*abstracte activiteit-namen*) en *4B* (*onlogische activiteit-namen*) aan dat er geen significant verschil is tussen de begrijpbaarheid van beide procesmodellen. Om meer zicht te krijgen op de impact van de *onlogische activiteit-namen* werden de scores van de participanten bij *procesmodel 4B* individueel bekeken. Op deze manier kan er gecontroleerd worden of er bepaalde participanten waren die wel moeilijkheden ondervonden met de *onlogische activiteit-namen* en daardoor veel fouten hebben gemaakt bij de vragen over *procesmodel 4B*. Hieruit bleek dat er geen enkele participant is die meer dan 1 fout extra maakte bij *procesmodel 4B* t.o.v. bij *procesmodel 3*. Bovendien is er maar één participant die slechts een score van 1 behaald op de *understandability task effectiveness* bij *model 4B* en deze participant had ook slechts een 2 bij *model 3*. Voor de rest zijn er geen participanten die het beduidend moeilijk hadden met *procesmodel 4B*. Er kan dus opnieuw afgeleid worden dat het gebruik van *onlogische activiteit-namen* weinig impact heeft op de begrijpbaarheid.

Om de resultaten verder te analyseren werd er bovendien opnieuw gekeken naar de verschillen tussen participanten met andere voorkennis (kennis van proces modelleren of BPMN). Voor de *understandability task effectiveness* zijn er geen opvallende verschillen. Bij *model 3* (*abstracte activiteit-namen*) en *4B* (*onlogische activiteit-namen*) zijn de medianen van de verschillende kennisgroepen heel erg gelijkend. Bij *model 4A* (*logische activiteit-namen*) liggen de medianen net iets hoger, maar er is geen opvallend verschil bij één van de kennisgroepen. Ook voor de *mentale inspanning* zijn er geen opvallende verschillen zichtbaar tussen de verschillende kennisgroepen. Bij *model 4A* zijn er meer participanten die een lage *mentale inspanning* aangaven.

Als laatste werd er per vraag ook gekeken naar het percentage van participanten dat een juist antwoord gaf over de verschillende modellen heen. Over het algemeen is dit percentage het hoogste bij *procesmodel 4A*. Enkel de percentages rond vraag 6 (*Q3-6* en *Q4A-6/Q4B-6*) zijn hierbij opvallend. Bij deze vraag is het percentage van juiste antwoorden het hoogste bij *procesmodel 4B* en het laagste bij *procesmodel 4A*. Om dit te verklaren werden allereerst de statistieken van deze vraag gedetailleerder bekeken. Van de participanten die *procesmodel 3* en *4A* kregen, gaven 3 participanten het foute antwoord op *Q3-6* en 6 participanten het foute antwoord op *Q4A-6*. Van de participanten die *procesmodel 3* en *4B* kregen, gaven 4 participanten het foute antwoord op *Q3-6* en gaf 1 participant het foute antwoord *Q4B-6*. Het is dus niet zo dat de participanten die vraag *Q3-6* fout beantwoorden toevallig allemaal *procesmodel 4A* kregen en ook hier *Q4A-6* fout beantwoorden. Vervolgens werd ook de inhoud van de vraag gedetailleerder bekeken. De participanten moesten hier aangeven hoe vaak een bepaalde activiteit maximaal voor kan komen (zie *Bijlage K, L* en *M*). Deze activiteit staat bij het onlogische proces op dezelfde plaats als bij het logische proces, waardoor er bij deze vraag geen noemenswaardig verschil is tussen *procesmodel 4A* en *procesmodel 4B* (zie *Bijlage L* en *M*). Dit zou een logische verklaring kunnen zijn voor het hogere percentage aan juiste antwoorden bij *procesmodel 4B*. Voor zover dit onderzoek gaat, kan er echter geen logische verklaring worden gegeven aan de andere percentages van juiste antwoorden bij deze vraag. Het gaat bovendien over percentages die redelijk dicht bij elkaar liggen: 89% bij *procesmodel 3*, 82% bij *procesmodel 4A* en 97% bij *procesmodel 4B*. T.o.v. het aantal participanten is dit verschil in principe verwaarloosbaar. Wanneer er per vraag gekeken werd naar de aangegeven *mentale inspanning* bij de verschillende procesmodellen kunnen er geen opvallende verschillen worden waargenomen.

Alle participanten hebben eerst *model 3* gekregen en vervolgens *model 4A* of *4B*. Dit heeft mogelijks een invloed gehad op de resultaten. Langs een kant zou het kunnen dat de participanten hebben geleerd uit *model 3* en hierdoor beter scoren op *model 4*. Langs de andere kant weten de participanten niet hoe goed ze gepresteerd hebben op *procesmodel 3* vooraleer ze *procesmodel 4* testen. Als de participanten beter of slechter scoren op *model 4* ligt dat hierdoor hoogstwaarschijnlijk aan *model 4* zelf, aangezien de participanten niet echt hebben kunnen leren uit de fouten die ze hebben gemaakt in *procesmodel 3*.

5.2.4 Kennis van participanten

In de literatuur wordt meerdere keren aangehaald en aangetoond dat persoonlijke factoren, zoals de kennis van proces modelleren, een grotere impact hebben op de begrijpbaarheid dan model gerelateerde factoren [36]. Mendling et al. toonden bovendien aan dat theoretische ervaring met proces modelleren vaak meer impact heeft dan praktische ervaring [25]. Als extra element werd er in dit onderzoek getracht de resultaten van deze voorgaande onderzoeken te valideren. Wanneer er gekeken werd naar de totale *understandability task effectiveness* per participant over de verschillende procesmodellen heen (score op 24), kan a.d.h.v.

de mediaan worden waargenomen dat de participanten zonder kennis van proces modelleren het laagste scoorde met een mediaan van 15. De participanten met praktische kennis van proces modelleren scoorden het hoogste met een mediaan van 21. En de participanten met theoretische kennis deden het met een mediaan van 20 juist iets minder goed. Om na te gaan of deze verschillen significant zijn werden er bovendien ook statistische testen uitgeoefend. Allereerst werd er a.d.h.v. de *Shapiro-Wilk test* gekeken of de data normaal verdeeld is. Aangezien enkel de data van de participanten met theoretische kennis normaal verdeeld is, werd er voor de vergelijking van de verschillende kennisgroepen gebruik gemaakt van een niet-parametrische test. Bovendien gaat het over de vergelijking van 3 onafhankelijke groepen in plaats van 2 onafhankelijke groepen. Daarom werd er gebruik gemaakt van de *Kruskal-Wallis One-Way ANOVA test* i.p.v. de *Mann-Whitney U-test* [3, 18]. Deze test toonde vervolgens aan dat er een significant verschil is tussen de verschillende groepen participanten op vlak van kennis van proces modelleren. Ook wanneer er gekeken werd naar de *mentale inspanning* worden precies dezelfde resultaten bekomen. Hieruit kan geconcludeerd worden dat ook in dit onderzoek de kennis van participanten een grote impact heeft op de begrijpbaarheid van procesmodellen [36]. Afgezien daarvan is het wel zo dat in dit onderzoek de participanten met praktische kennis beter scoorde dan de participanten met theoretische kennis. Dit ligt niet in lijn met de resultaten van het onderzoek van Mendling et al. [25].

6 Beperkingen en toekomstig onderzoek

Er zijn een aantal beperkingen die in rekening gebracht moeten worden bij het interpreteren van de resultaten van dit onderzoek. Hieronder worden deze beperkingen kort toegelicht en worden er enkele suggesties gegeven voor toekomstig onderzoek.

Ten eerste, ondanks het meegegeven advies om de enquête in te vullen op een pc of laptop, hebben enkele participanten dit niet gedaan. Van de 66 participanten vulden 5 participanten de enquête in op een gsm en vulde 1 participant de enquête in op nog een ander toestel (mogelijk een ipad). Aangezien een gsm vaak een kleiner scherm heeft dan een pc of laptop moet de lezer meer scrollen doorheen de procesmodellen en zijn aandacht verdelen over meerdere weergaven. Dit kan een negatieve effect hebben gehad op de begrijpbaarheid [16, 33, 40, 41]. In dit onderzoek werd geen rekening gehouden met dit effect. Toekomstig onderzoek zou het gebruik van een pc kunnen verplichten of rekening kunnen houden met de impact van het gebruik van een gsm voor het invullen van de enquête.

Ten tweede is het belangrijk om rekening te houden met het feit dat elke participant *procesmodel 3* (*abstracte-activiteit namen*) heeft getest en vervolgens ook vragen kreeg over *procesmodel 4A* (*logische activiteit-namen*) of *4B* (*onlogische activiteit-namen*). Het beantwoorden van de vragen bij *procesmodel 3* kan hierdoor een impact hebben gehad op het beantwoorden van vragen bij *procesmodel 4A* of *4B*, zeker aangezien de vragen over de modellen *syntactisch* hetzelfde waren. Deze beperking werd verminderd door de modellen niet

rechtstreeks na elkaar weer te geven en door de vragen bij *procesmodel 4A* en *4B* op een andere volgorde te zetten. Bovendien kregen de participanten geen verbetering of score na het invullen van de vragen bij *procesmodel 3*, waardoor de participanten niet echt konden leren uit hun gemaakte fouten. In de toekomst zou het echter beter zijn om te werken met allemaal onafhankelijke data en *procesmodel 3*, *4A* en *4B* te laten invullen door drie verschillende groepen participanten.

Als laatste moet er ook verwezen worden naar de complexe relatie tussen de verschillende factoren van de begripbaarheid [6]. Zoals eerder aangehaald vertonen voorgaande onderzoeken vaak tegenstrijdige resultaten die te danken zijn aan deze complexe relatie [6]. Naarmate deze factoren, bv. de grootte van het procesmodel, de modeleringsnotatie, het presentatie medium, het aantal beslissingen, etc., verschillen, kunnen ook de resultaten van de onderzoeken anders zijn [6]. Als er in dit onderzoek andere modellen waren gebruikt voor de enquête, waren er wellicht andere resultaten gevonden. Ook andere vragen hadden misschien tot andere conclusies geleiden. Bovendien is het ook niet zeker of de vragen die gebruikt werden in de enquête een juist beeld geven over de begripbaarheid en of alle aspecten van begripbaarheid hierdoor worden afgedekt.

7 Conclusie

Procesmodellen zijn een veelgebruikt hulpmiddel bij het implementeren, beheren en verbeteren van processen en worden bovendien gebruikt om de communicatie tussen de verschillende stakeholders te vergemakkelijken. Door deze verschillende stakeholders is het van uiterst belang om rekening te houden met de begripbaarheid van procesmodellen. Om meer inzicht te krijgen in deze begripbaarheid werd er in deze paper onderzocht hoe verschillende representaties van *modulariteit*, *condities* en *activiteit-namen* de begripbaarheid van procesmodellen in BPMN beïnvloeden.

A.d.h.v. een empirisch onderzoek rond 7 verschillende procesmodellen werden er enkele bevindingen vastgelegd die een meerwaarde zouden kunnen zijn bij het ontwerpen van procesmodellen. Allereerst werd er ontdekt dat bij het weergeven van *modulariteit* er het beste gebruik kan worden gemaakt van *expanded sub-processen* i.p.v. *collapsed sub-processen*. Hierbij werd er bovendien geconcludeerd dat het voor lezers van procesmodellen met *collapsed sub-processen* vooral moeilijker is wanneer er vragen worden gesteld waarbij verschillende elementen van binnen en buiten de sub-processen samen bekeken moesten worden. Vervolgens werd er ook ontdekt dat de begripbaarheid van het gebruik van *XOR-gateways* ongeveer even groot is als de begripbaarheid van *intermediate error events*. Bovendien werd er geconcludeerd dat het gebruik van *onlogische activiteit-namen* geen impact heeft op de begripbaarheid van procesmodellen, maar dat lezers het wel makkelijker hebben met het begrijpen van een procesmodel met *logische activiteit-namen* t.o.v. een procesmodel met *abstracte activiteit-namen*. Additiekoneel werd er ook ontdekt dat de kennis van de lezer een grote impact heeft op de begripbaarheid. Hoe groter de kennis van de lezer, hoe groter de *under-*

standability task effectiveness en hoe lager de vereisten *mentale inspanning* voor het begrijpen van het procesmodel. Theoretische training voor de verschillende stakeholders, die moeten werken met procesmodellen, zou daarom veel voordelen kunnen opleveren voor een onderneming. In tegenstelling tot het voorgaande onderzoek van Mendling et al. werd er in dit onderzoek wel ondervonden dat participanten met praktische kennis nog beter scoorde op het begrijpen van de procesmodellen dan participanten met enkel theoretische kennis. A.d.h.v. deze bevindingen kan deze paper bijdragen aan het ontwerpen van begrijpbare procesmodellen.

In grote lijnen bevat dit onderzoek twee belangrijke beperkingen die een richting kunnen aangeven voor toekomstige onderzoeken. Ten eerste werden de testen rond de begrijpbaarheid van *activiteit-namen* niet volledig onafhankelijk van elkaar uitgevoerd. De participanten kregen namelijk twee keer dezelfde vragen over hetzelfde procesmodel met andere *activiteit-namen*. Ten tweede zijn de relaties tussen de verschillende factoren van begrijpbaarheid redelijk complex, waardoor er wellicht andere resultaten zouden worden gevonden als er gebruikt werd gemaakt van andere procesmodellen en/of andere vragen om de begrijpbaarheid te operationaliseren.

Dankwoord

Allereerst zou ik graag mijn promotor dr. Frank Vanhoenshoven willen bedanken voor zijn begeleiding, waardevolle suggesties en tijd in de voorbije zomermaanden. Als tweede zou ik graag de participanten van mijn onderzoek willen bedanken, aangezien hun bijdrage een grote invloed heeft gehad op mijn onderzoek. Vervolgens zou ik graag Mathijs Stevens willen bedanken voor het testen van mijn enquête en voor het nalezen van mijn masterproef. Ten slotte zou ik graag mijn ouders willen bedanken voor hun steun en voor de kansen die ze mij gegeven hebben om mijn studies af te kunnen ronden aan de UHasselt.

Referenties

1. Aldin, L., de Cesare, S.: A comparative analysis of business process modelling techniques. UKAIS (2009)
2. Becker, J., Rosemann, M., von Uthmann, C.: Guidelines of Business Process Modeling. In: van der Aalst, W., Desel, J., Oberweis, A. (eds.) Business Process Management, pp. 30–49. Springer Berlin Heidelberg, Berlin, Heidelberg (2000)
3. Bettany-Saltikov, J., Whittaker, V.J.: Selecting the most appropriate inferential statistical test for your quantitative research study. *Journal of Clinical Nursing* (11-12), 1520–1531 (2014)
4. Corradini, F., Ferrari, A., Fornari, F., Gnesi, S., Polini, A., Re, B., Spagnolo, G.O.: A Guidelines framework for understandable BPMN models. *Data & Knowledge Engineering* pp. 129–154 (2018)
5. De A. Rodrigues, R., De O. Barros, M., Revoredo, K., Azevedo, L.G., Leopold, H.: An Experiment on Process Model Understandability Using Textual Work Instructions and BPMN Models. In: 2015 29th Brazilian Symposium on Software Engineering. pp. 41–50 (2015)

6. Dikici, A., Turetken, O., Demirors, O.: Factors influencing the understandability of process models: A systematic literature review. *Information and Software Technology* pp. 112–129 (2018)
7. Dumas, M., La Rosa, M., Mendling, J., Reijers, H.A.: *Fundamentals of Business Process Management*. Springer Berlin Heidelberg, Berlin, Heidelberg (2013)
8. Figl, K., Strembeck, M.: Findings from an Experiment on Flow Direction of Business Process Models. In: *EMISA* (2015)
9. Figl, K., Laue, R.: Cognitive Complexity in Business Process Modeling. pp. 452–466 (2011)
10. Friedrich, F., Holschke, O., Recker, J., Kohlborn, T., Breitzkreuz, D., Reijers, H.: Automated Generation of Business Process Models from Natural Language Input (2010)
11. Gabryelczyk, R., Jurczuk, A.: Does Experience Matter? Factors Affecting the Understandability of the Business Process Modelling Notation. *Procedia Engineering* pp. 198–205 (2017)
12. Harel, D., Rumpe, B.: Meaningful modeling: What’s the semantics of ”semantics”? *Computer* **37**, 64–72 (2004)
13. Jošt, G., Huber, J., Polančič, G., Hericko, M.: An empirical investigation of intuitive understandability of process diagrams. *Computer Standards & Interfaces* (2016)
14. Leopold, H., Smirnov, S., Mendling, J.: On the refactoring of activity labels in business process models. *Information Systems* (5), 443–459 (2012)
15. Lieben, J., Jouck, T., Depaire, B., Jans, M.: An Improved Way for Measuring Simplicity During Process Discovery: 14th International Workshop, EOMAS 2018, Held at CAiSE 2018, Tallinn, Estonia, June 11–12, 2018, Selected Papers. pp. 49–62 (2018)
16. Lübke, D., Ahrens, M., Schneider, K.: Influence of diagram layout and scrolling on understandability of BPMN processes: an eye tracking experiment with BPMN diagrams. *Information Technology and Management* (2), 99–131 (2021)
17. Maes, A., Poels, G.: Evaluating Quality of Conceptual Models Based on User Perceptions. pp. 54–67 (2006)
18. McCrum-Gardner, E.: Which is the correct statistical test to use? *British Journal of Oral and Maxillofacial Surgery* (1), 38–41 (2008)
19. Mendling, J., Reijers, H.A., Recker, J.: Activity labeling in process modeling: Empirical insights and recommendations. *Information Systems* (4), 467–482 (2010)
20. Mendling, J., Neumann, G., van der Aalst, W.: Understanding the Occurrence of Errors in Process Models Based on Metrics. In: Meersman, R., Tari, Z. (eds.) *On the Move to Meaningful Internet Systems 2007: CoopIS, DOA, ODBASE, GADA, and IS*. pp. 113–130. *Lecture Notes in Computer Science*, Springer (2007)
21. Mendling, J., Recker, J., Reijers, H.: On the Usage of Labels and Icons in Business Process Modeling. *International Journal of Information System Modeling and Design* (2010)
22. Mendling, J., Reijers, H., Aalst, W.: Seven Process Modeling Guidelines (7PMG). *Information and Software Technology* pp. 127–136 (2010)
23. Mendling, J., Reijers, H., Cardoso, J.: What Makes Process Models Understandable? (2007)
24. Mendling, J., Strembeck, M.: Influence Factors of Understanding Business Process Models. pp. 142–153 (2008)
25. Mendling, J., Strembeck, M., Recker, J.: Factors of process model comprehension—Findings from a series of experiments. *Decision Support Systems* pp. 195–206 (2012)

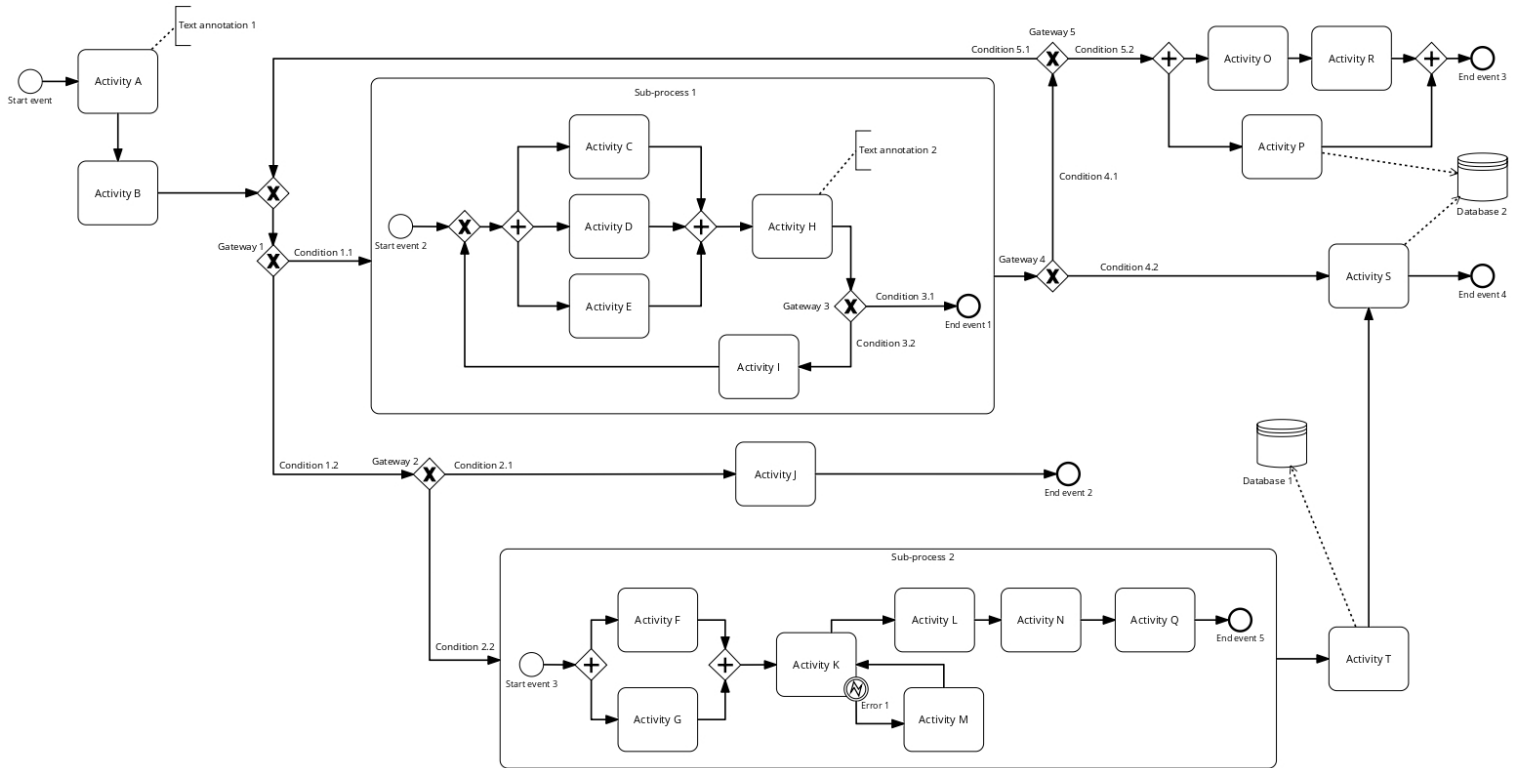
26. Mishra, P., Pandey, C.M., Singh, U., Gupta, A., Sahu, C., Keshri, A.: Descriptive Statistics and Normality Tests for Statistical Data. *Annals of Cardiac Anaesthesia* (1), 67–72 (2019)
27. Moreno Montes de Oca, I., Snoeck, M.: Pragmatic Guidelines for Business Process Modeling. *SSRN Electronic Journal* (2014)
28. Ottensooser, A., Fekete, A., Reijers, H., Mendling, J., Menictas, C.: Making sense of business process descriptions: An experimental comparison of graphical and textual notations. *Journal of Systems and Software - JSS* (2012)
29. Paas, F., van Merriënboer, J.J.G.: Cognitive-Load Theory: Methods to Manage Working Memory Load in the Learning of Complex Tasks. *Current Directions in Psychological Science* (4), 394–398 (2020)
30. Patig, S., Casanova-Brito, V., Vögeli, B.: IT Requirements of Business Process Management in Practice - An Empirical Study. In: *BPM* (2010)
31. Purchase, H.: Which aesthetic has the greatest effect on human understanding? In: Goos, G., Hartmanis, J., van Leeuwen, J., DiBattista, G. (eds.) *Graph Drawing*, pp. 248–261. Springer Berlin Heidelberg (1997)
32. Recker, J.: Empirical Investigation of the Usefulness of Gateway Constructs in Process Models. *European Journal of Information Systems* p. forthcoming (2013)
33. Reijers, H.A., Mendling, J., Dijkman, R.M.: Human and automatic modularizations of process models to enhance their comprehension. *Information Systems* (5), 881–897 (2011)
34. Reijers, H., Mendling, J., Dijkman, R.: On the usefulness of subprocesses in business process models. *BPM reports*, BPMcenter. org (2010)
35. Reijers, H., Mendling, J.: Modularity in Process Models: Review and Effects. In: Dumas, M., Reichert, M., Shan, M.C. (eds.) *Business Process Management*. pp. 20–35. *Lecture Notes in Computer Science*, Springer (2008)
36. Reijers, H.A., Mendling, J.: A Study Into the Factors That Influence the Understandability of Business Process Models. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* (3), 449–462 (2011)
37. Rosing, M.V.: *The complete business process handbook: body of knowledge from process modeling to bpm*. Elsevier, Waltham, MA, 1st edition edn. (2014)
38. Sharp, A., Mcdermott, P.: *Workflow Modeling: Tools for Process Improvement and Application Development* (2001)
39. Sánchez-González, L., Ruiz, F., García, F., Cardoso, J.: Towards thresholds of control flow complexity measures for BPMN models. In: *Proceedings of the 2011 ACM Symposium on Applied Computing*. pp. 1445–1450. SAC '11, Association for Computing Machinery (2011)
40. Turetken, O., Dikici, A., Vanderfeesten, I., Rompen, T., Demirors, O.: The Influence of Using Collapsed Sub-processes and Groups on the Understandability of Business Process Models. *Business & Information Systems Engineering* (2), 121–141 (2020)
41. Turetken, O., Rompen, T., Vanderfeesten, I., Dikici, A., Moll, J.: The Effect of Modularity Representation and Presentation Medium on the Understandability of Business Process Models in BPMN. pp. 289–307 (2016)
42. Weber, B., Reichert, M.: Refactoring Process Models in Large Process Repositories. pp. 124–139 (Jun 2008)
43. White, S.A., Miers, D.: *BPMN modeling and reference guide: understanding and using BPMN ; develop rigorous yet understandable graphical representations of business processes*. Future Strategies Inc (2008)
44. Yadira, V., Silveira, D., Albuquerque, M., Araújo, J.: Verifying BPMN understandability with novice business managers. pp. 94–101 (2018)

45. Zugal, S., Pinggera, J., Reijers, H., Reichert, M., Weber, B.: Making the case for measuring mental effort. In: Proceedings of the Second Edition of the International Workshop on Experiences and Empirical Studies in Software Modelling - EESSMod '12. p. 1. ACM Press (2012)

8 Bijlagen

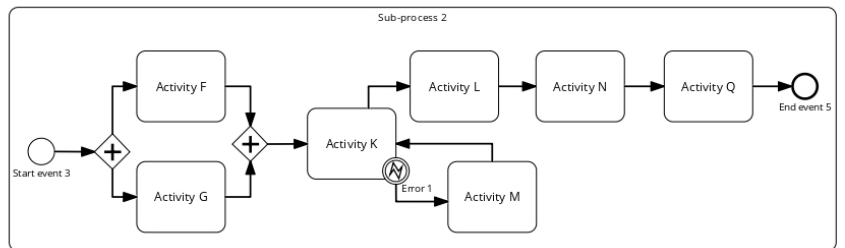
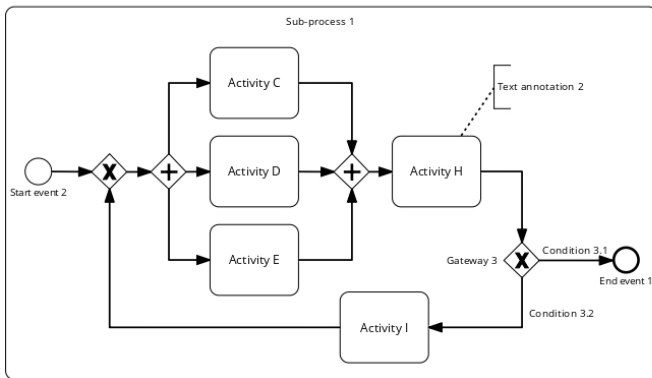
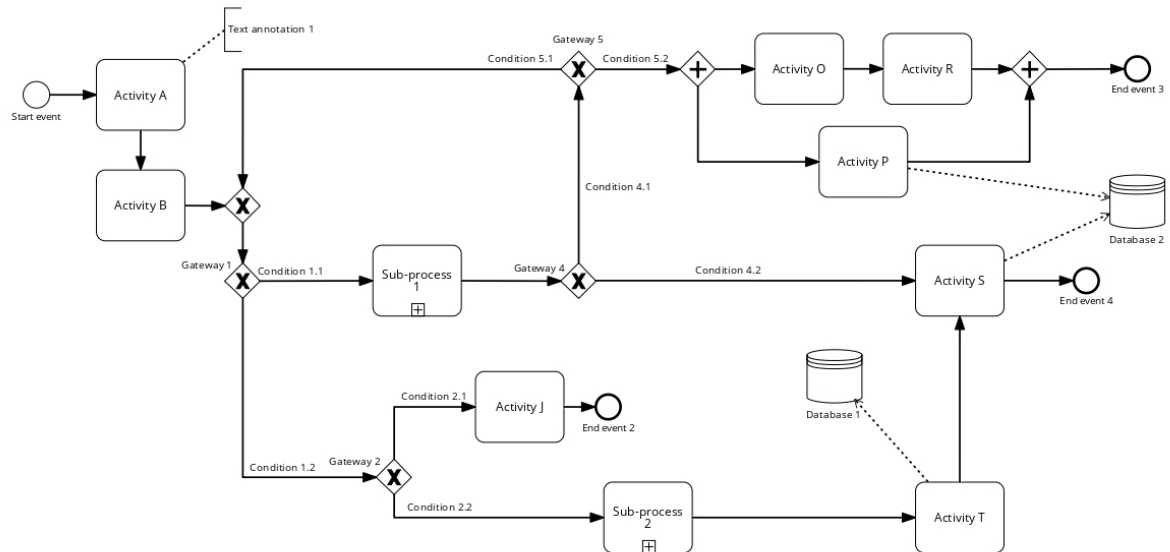
A

Procesmodel 1A (expanded sub-processen)



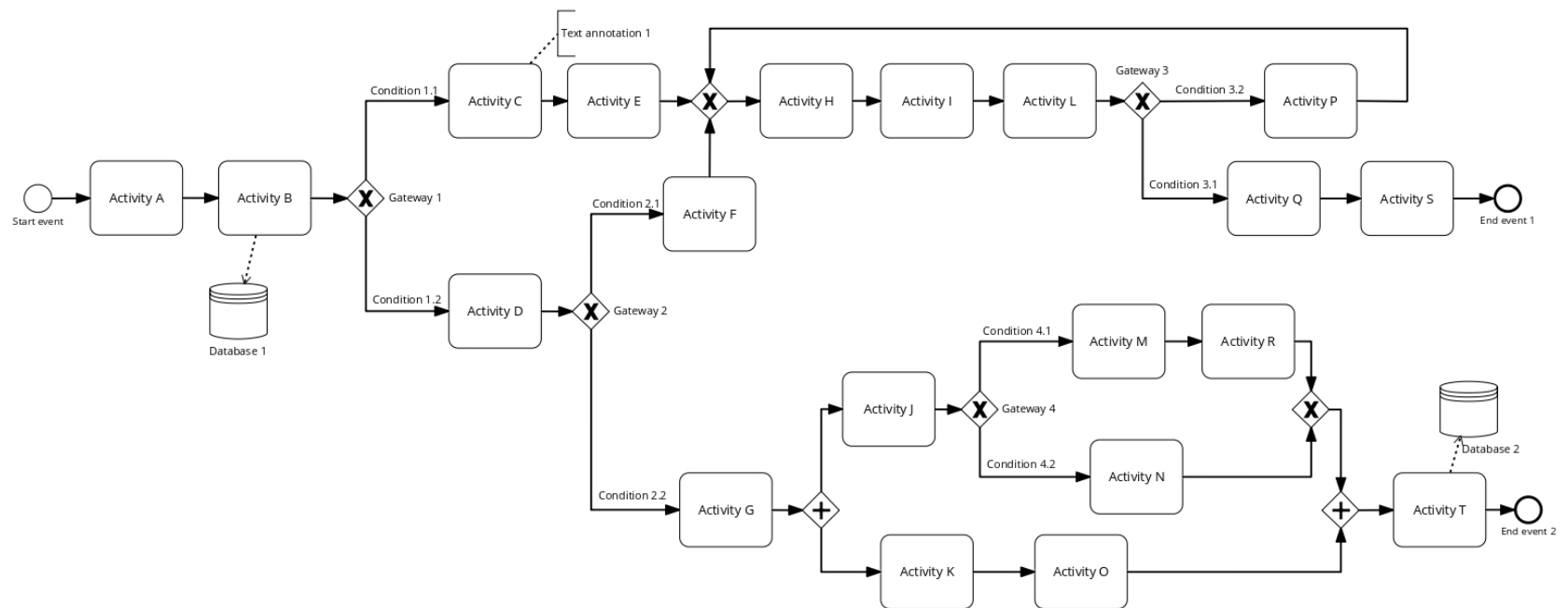
B

Procesmodel 1B (collapsed sub-processen)



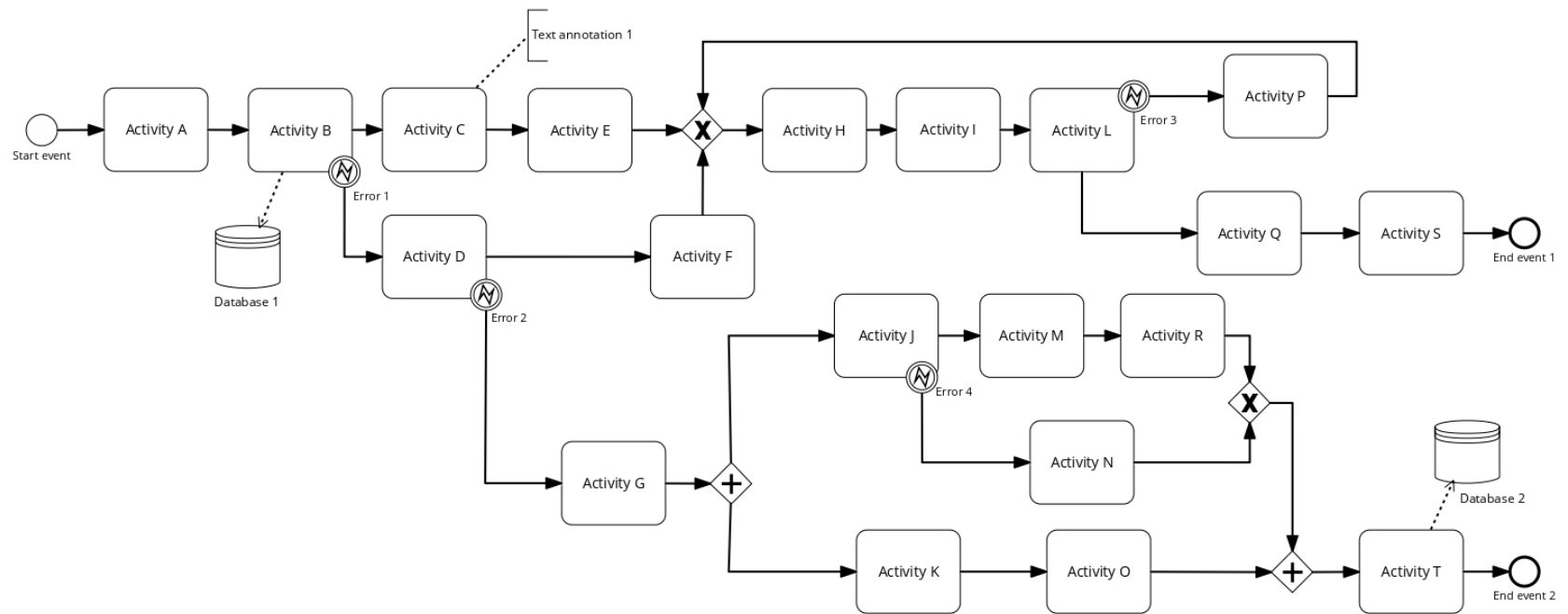
C

Procesmodel 2A (XOR-gateways)



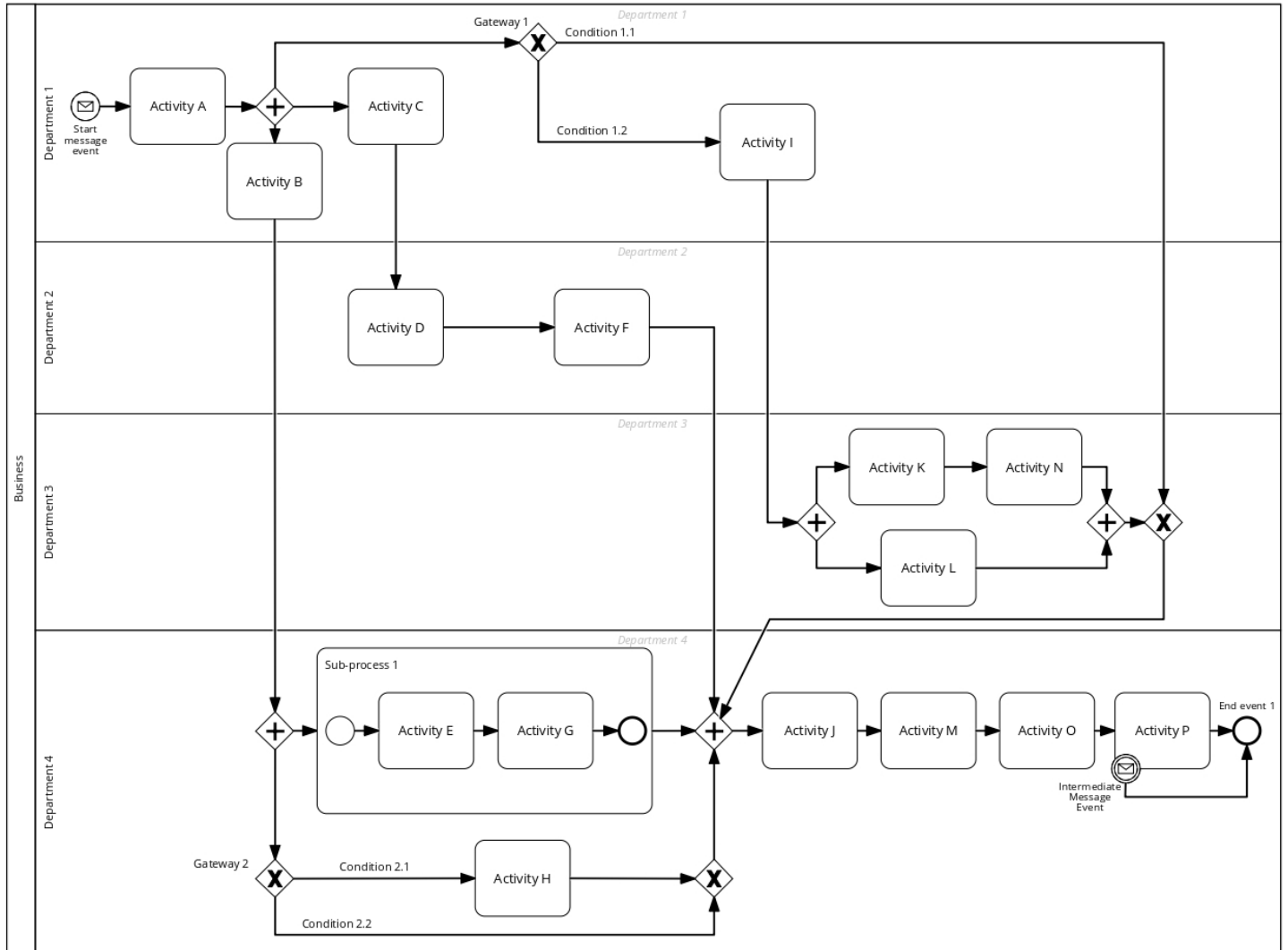
D

Procesmodel 2B (intermediate error events)



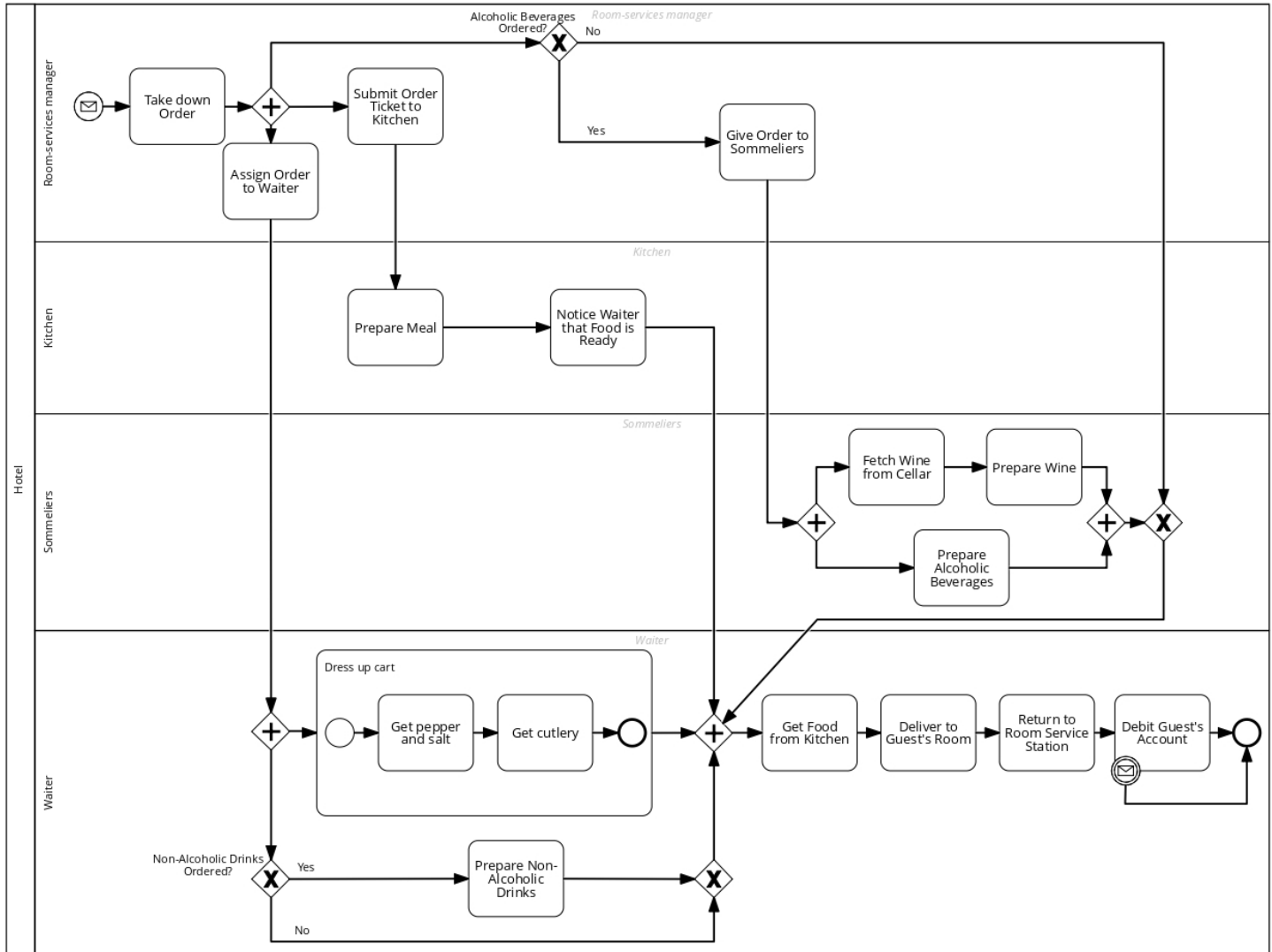
E

Procesmodel 3 (abstracte activiteit-namen)



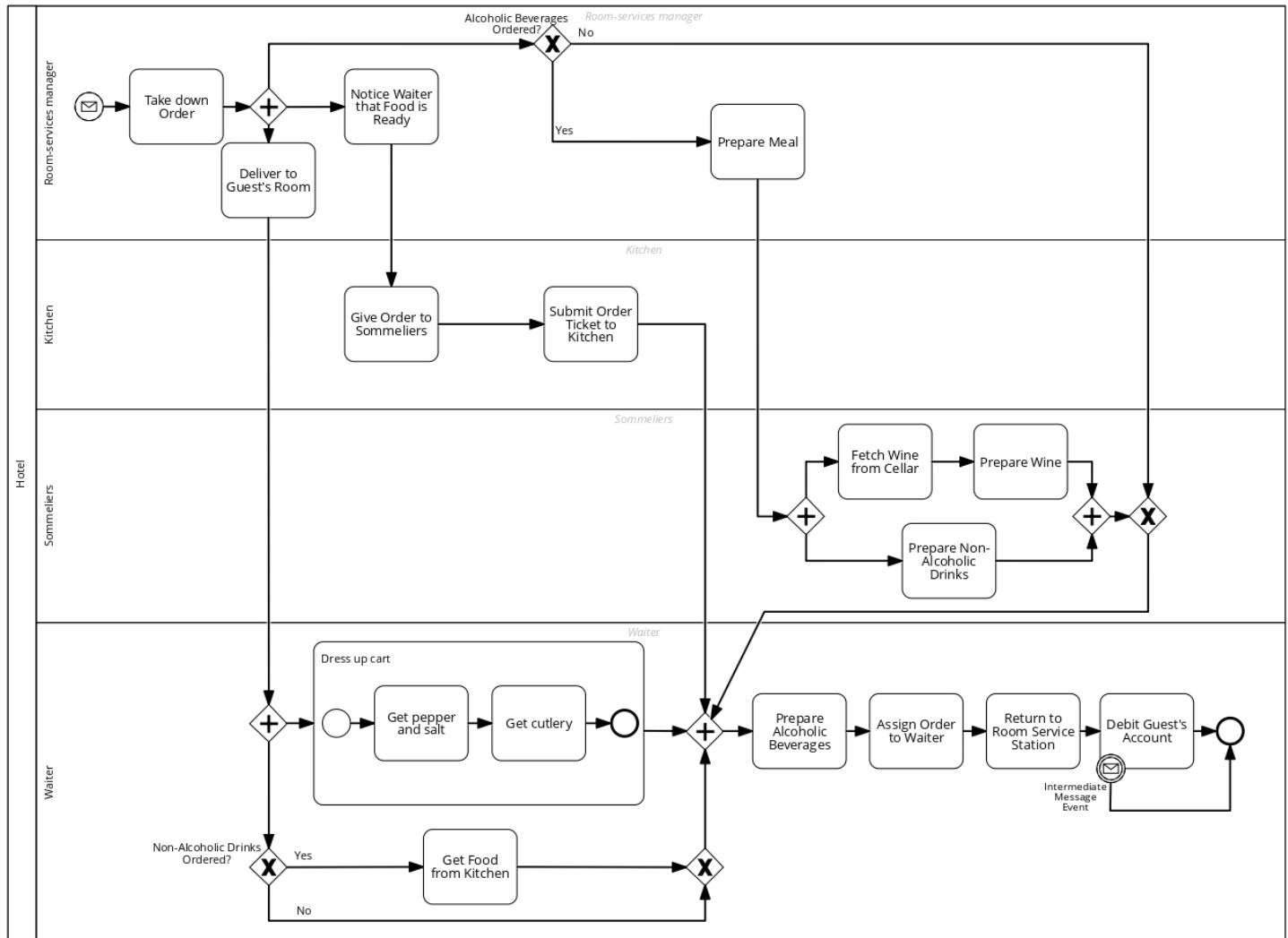
F

Procesmodel 4A (Logische activiteit-namen)



G

Procesmodel 4B (Onlogisch activiteit-namen)



H

Inleidende vragen

	Vraag	Antwoordopties
Q0-1	Are you a student?	Yes - No
Q0-2	What level of education do you have? (Please select the highest educational qualification you have reached until now)	No diploma - High school degree - Bachelor's degree - Master's degree
Q0-3	How would you rate your prior knowledge in the modeling language BPMN?	I have no prior knowledge of process modeling - I have theoretical knowledge of process modeling - I have practical knowledge of process modeling
Q0-4	How would you rate your prior knowledge in the modeling language BPMN?	I never heard of it - ... - ... - ... - I'm an expert
Q0-5	Which device are you using to fill out this survey?	Mobile phone - PC/Laptop - Other

I

Begrijpbaarheidsvragen procesmodel 1A en 1B

	Vraag	Antwoordopties (vetgedrukte is juist antwoord)
Q1A-1 Q1B-1	Activity H can be performed at maximum:	One time - Two times - Three times - Infinite times - I don't know
Q1A-2 Q1B-2	The next activity, after activity Q, is activity:	Activity J - Activity S - Activity T - Activity O - I don't know
Q1A-3 Q1B-3	Which activity can't directly follow activity H? (By 'directly follow' we mean that it is possible to carry out a certain activity right after activity H)	Activity J - Activity T - Activity C - Activity F - I don't know
Q1A-4 Q1B-4	How many different direct predecessors (forerunners) does activity S have? (By 'direct predecessors' we mean the activities that happen directly before activity S)	One - Two - Three - Four - I don't know
Q1A-5 Q1B-5	Does activity H always occur?	Yes - No - I don't know
Q1A-6 Q1B-6	What is the minimum number of activities needed to get into end event 4?	[1-25] - I don't know (Antwoord: 7)

J

Begrijpbaarheidsvragen procesmodel 2A en 2B

	Vraag	Antwoordopties (vetgedrukte is juist antwoord)
Q2A-1 Q2B-1	Does activity D always occur?	Yes - No - I don't know
Q2A-2 Q2B-2	Activity F can be performed at maximum:	One time - Two times - Three times - Infinite times - I don't know
Q2A-3 Q2B-3	Look at activity N and activity R. Which of the following statements is correct:	Activity N can happen before activity R - Activity N always happens before activity R - Activity N and activity R can't happen both - I don't know
Q2A-4 Q2B-4	Activity H can be performed at maximum:	One time - Two times - Three times - Infinite times - I don't know
Q2A-5 Q2B-5	Can activity S happen before activity P happens?	Yes - No - I don't know
Q2A-6 Q2B-6	Does activity E always occur?	Yes - No - I don't know

K

Begrijpbaarheidsvragen procesmodel 3

	Vraag	Antwoordopties (vetgedrukte is juist antwoord)
Q3-1	Does sub-process 1 always happen?	Yes - No - I don't know
Q3-2	What is the minimum number of activities needed to get to activity J? (Activity J not included)	[1-25] - I don't know (Antwoord: 7)
Q3-3	What activity does department 3 not perform?	Activity K - Activity L - Activity J - Activity N - I don't know
Q3-4	Which of the following statements is correct?	Activity M happens before activity J - Activity I does only happen when condition 1.2 is true - Activity E happens after activity P - I don't know
Q3-5	Which of these activities happens first?	Activity J - Activity B - Activity H - I don't know
Q3-6	How many times does activity K happen at most?	One time - Two times - Three times - Infinite times - I don't know

L

Begrijpbaarheidsvragen procesmodel 4A

	Vraag	Antwoordopties (vetgedrukte is juist antwoord)
Q4A-1	Does the sub-process 'Dress up cart' always happen?	Yes - No - I don't know
Q4A-2	What is the minimum number of activities needed to get to the activity 'Get Food from Kitchen'? (Activity 'Get Food from Kitchen' not included)	[1-25] - I don't know (Antwoord: 7)
Q4A-3	What activity do the sommeliers not perform?	Fetch Wine from Cellar - Prepare Alcoholic Beverages - Get Food from Kitchen - Prepare Wine - I don't know
Q4A-4	Which of the following statements is correct?	The activity 'Deliver to Guest's Room' happens before the activity 'Get Food from Kitchen' - The activity 'Give order to Sommeliers' does only happen when the customer ordered alcoholic beverages - The activity 'Get pepper and salt' happens after the activity 'Debit Guest's Account' - I don't know
Q4A-5	Which of these activities happens first?	Get Food from Kitchen - Assign Order to Waiter - Prepare Non-Alcoholic Drinks - I don't know
Q4A-6	How many times does activity 'Fetch Wine from Cellar' happen at most?	One time - Two times - Three times - Infinite times - I don't know

M

Begrijpbaarheidsvragen procesmodel 4B

	Vraag	Antwoordopties (vetgedrukte is juist antwoord)
Q4B-1	Does the sub-process 'Dress up cart' always happen?	Yes - No - I don't know
Q4B-2	What is the minimum number of activities needed to get to the activity 'Prepare Alcoholic Beverages'? (Activity 'Prepare Alcoholic Beverages' not included)	[1-25] - I don't know (Antwoord: 7)
Q4B-3	What activity do the sommeliers not perform?	Fetch Wine from Cellar - Prepare Non-Alcoholic Drinks - Prepare Alcoholic Beverages - Prepare Wine - I don't know
Q4B-4	Which of the following statements is correct?	The activity 'Assign Order to Waiter' happens before the activity 'Prepare Alcoholic Beverages' - The activity 'Prepare Meal' does only happen when the customer ordered alcoholic beverages - The activity 'Get pepper and salt' happens after the activity 'Debit Guest's Account' - I don't know
Q4B-5	Which of these activities happens first?	Prepare Alcoholic Beverages - Deliver to Guest's Room - Get Food from Kitchen - I don't know
Q4B-6	How many times does activity 'Fetch Wine from Cellar' happen at most?	One time - Two times - Three times - Infinite times - I don't know