# Applying structure to benefits of Automated Testing

Frederik Ørbæk Hendriksen and Christian Kirk Damsgaard

Department of Computer Science, Aalborg University,

Selma Lagerlöfs Vej 300, 9220 Aalborg Øst, Denmark

{fhendr14, cdth14}@student.aau.dk

## SUMMARY

Automated testing (AT) has been researched for three decades regarding its benefits over manual testing (MT). It may be observed to have grown in implementation and considerations of software companies' visions and strategies. AT may be defined as a set of instructions used to verify the functionality it has been built to test. AT, compared to MT, may be observed to have multiple benefits but needs the right conditions to realise its potential. Therefore, analysing AT as an IT investment is essential before implementing it properly into an organisation, as it may be time-consuming and cost-heavy. To explore how organisations incorporate and disseminate AT, we conducted an Action Research study in collaboration with Elbek & Vejrup, a Danish software- and consultancy organisation. We conducted two interventions applying an adapted business case model (BC model), where we first introduced unstructured and later structured benefits to challenging their conception of AT benefits. The process of writing this article started with us researching the concept of AT after meeting with Elbek & Vejrup. We were tasked with a presentation concerning AT to their consultants and product owners, which came to revolve around benefits.

During this presentation and the following panel debate, we learned that when the benefits of AT are disseminated unstructured, following discussions will also be unstructured, only revolving around how much it costs to implement and how much time it takes. Furthermore, we observed that there were no comments or insights regarding the risks of AT, which made us reconsider which aspects to incorporate into our utilisation of the Business case model for IT investments. Two aspects we came about during our research in the existing AT benefits literature was an extended focus on limitations and conditions rather than risks. We, therefore, included limitations and conditions when structuring the benefits for us to explore if the second intervention would bring about further dissemination of AT when applying a structure to the benefits. For our second intervention, we had structured the benefits of AT, which we could observe in the existing literature. We disseminated the structured benefits using the benefits grid of the Business case model. We disseminated and discussed the structured benefits to product owners for interviews that spanned 1,5 hours each. We observed during the interviews that the product owner started to use the terminology of the benefits grid, disseminating each benefit by using the type of change and -valuation. We further observed that the product owners referred almost every benefit to either a limitation or a condition without us prompting them for these references. The product owners did not disseminate the limitations and conditions during the interviews. However, when they

reflected on benefits and mentioned conditions or limitations, we would challenge their experience, prompting them for elaboration. We analysed and reported the findings of the interventions and discussed them with the application of the BC model in the existing literature and practice. In the case of the dissemination of AT benefits, we observed a significant change in how the consultants and product owners articulated their answers when the structure of the Business case model's benefits grid was applied to the same benefits that had been presented and discussed earlier. Our Action Research approach impacted the organisation, which we collaborated with, and several product owners used the Business case model's benefits grid terminology to disseminate how beneficial implementation of a new function may be: especially the type of change. Our Action Research approach furthermore indicated how to consider and expand the existing literature on AT, bringing a nuanced and multidimensional structure to a topic that consists of separate lists for presenting findings regarding benefits.

# 1   Introduction

Testing within an IT company generally requires many resources to acquire a product with few or no errors successfully. Therefore, companies have begun implementing Automated Testing (AT) to lower the number of resources spent on testing [1]. AT has been researched since the 1990s due to an expansion in implementations of the concept [2]–[5]. Furthermore, the requirements come with potential benefits and limits that the IT organisation should be aware of, such as price, time, and quality.

Testing plays a central part in an IT organisation's way to improve on its product by doing manual testing (MT) and AT, but one of the issues with MT comes with the time invested into doing the test compared to AT. Therefore, companies are becoming convinced that if done correctly, AT is superior [1]. However, for organisations to achieve operational excellence or to get an edge over competing organisations, they must innovate; organisations must make investments [6]. Therefore, the benefits of AT are worth considering from an investment perspective and identifying and structuring benefits becomes essential to address the investment's feasibility. We believe we have observed a gap in the existing literature on AT, which may enhance evaluations of the feasibility regarding benefits and risks. To embrace the complexity of the technological investment, we believe the introduction of structurisation, and dissemination of benefits needs to be considered.

To our knowledge, the current models and methods for evaluating AT in an organisational perspective is based on listings of benefits and limitations [7]–[9], and an evaluation model based on a cost/benefit analysis [10]. We need to structure identified benefits within an organisation, where we suggest using the Business Case Model (BC model) [11]. The BC model is a framework for identifying and structuring benefits, measures and owners, and documentation representation [11]–[13]. We will explore these considerations while conducting an Action Research study [14]. To our knowledge, there are no studies of using the AR method with the BC

model regarding implementing AT. However, studies show the AR study as a competent tool to research the theoretical background concerning a topic and build upon it [14]. Therefore, with this AR study, we intend to identify and structure benefits by introducing unstructured and structured benefits to the company Elbek & Vejrup (EV) and explore how the BC model may contribute to the dissemination of benefits within EV. We will explore and research the proposed gap using the BC model in an AR study within the organisation based on the research question: *"How can we impact the dissemination of Automated testing at a software company by using a business case model?"*.

## 2 Background

### 2.1 Benefits of automated testing

Research regarding benefits and value have been studied over the last three decades, and we find the most essential and contemporary contribution to this topic to be Rafi et al. [7] and Karhu et al. [9], which may be supplemented with a systemic literature study, conducted by Lindholm [8]. Rafi et al. Presented a list of 9 benefits, which concerned B1) improved quality, B2) test coverage, B3) reduced time for testing, B4) reliability regarding repeating tests, B4) increased confidence regards to the system, B6) test-reusability, B7) reducing human effort, B8) reduction of cost and B9) increased fault detection [7].

Karhu et al. studied the impact of AT in different organisations, where they had six significant observations, which mention both the benefits and limitations of AT. Their first observation concerning cost reduction, improved quality and less time spent on testing while bringing about- and allocating costs. The second observation revealed that generic- and independent systems facilitate AT, while complex systems cause problems. Their third observation focused on uniform vs different systems concerning ease of implementation of AT. The fourth observation related to the extent of human involvement, stating that less involvement facilitates AT. The fifth observation concerned the span of lifecycle and rapid changes, stating that rapid changes may bring about obsolete tests. The last observation by Karhu et al. revolved around reusability and the importance of a long-term focus when investing in AT. As presented, Karhu et al. did not only focus on the benefits but also made observations regarding which parameters a system and development team should consider before investing in AT [9]. Rafi et al. also had their focus on the limitations of AT, where they presented seven limitations: L1) AT does not replace manual testing, L2) the means justifies the end, 3) fast-paced technological evolution makes AT challenging to maintain, L4) AT process requires time, L5) False expectations regarding cost-saving, L6) Test strategy the does not allow the realisation of the benefits of AT, L7) Lack of skilled staff [7]. These considerations of benefits and limitations have been utilised by Lindholm in his *"Economics of test automation"* [8], a systemic literature review which included the benefits and limitations of AT. Lindholm concluded that AT might bring both short- and long-term benefits in an organisational context, especially human involvement in testing, reduction of costs, and short lifecycles (release cycles). Lindholm

further elaborated that its parameters for measuring the realisation may be measured (with a positive result) within four years but may be limited by the organisation's culture regarding the investment into AT. Lindholm further concluded that in the specific context in which the benefits and limitations of AT were studied, seven out of Rafi et al.'s proposed nine benefits were believed to be realised by the practitioners, excluding B3) reduced time for testing and B8) reduction of cost. This conclusion makes us ponder the disagreement regarding Lindholm's findings concerning B3 and B8, concerning Karhu et al.'s first observation.

We believe that Rafi et al. has produced a list of benefits to consider. At the same time, Karhu et al. bring about empirical observations and a mode for considering parameters and conditions of systems for the benefits to be realisable. In contrast, Lindholm's systemic research has brought about an approach to further considering benefits in an organisational context.

## 2.2 Business case model for IT-investments

These considerations regarding analysing AT as an investment do not provide a structure or model of how benefits may be disseminated in an organisational context, which is why we would like to introduce the Business case model for IT investments (BC model). The BC model was developed by Ward et al. [12]. They developed this model on the foundation that a survey they conducted with over 100 organisations had issues identifying and quantifying benefits when considering investing in new IT. This model brought about a new model of building a business case, where the focus is on types of benefits, how their measures are identified, and calls for the identification of the owner of each benefit. In addition, the BC model has been introduced to adaptations. Nielsen and Persson adapted their use, excluding *quantifiable benefits* due to observations of struggles with this structurisation of benefits by practitioners[11], [13]. Nielsen and Persson adapted to BC model to cover three types of benefits while still utilising Ward et al.'s proposed types of change, see table 1.

|  |  | Type of change | | |
| --- | --- | --- | --- | --- |
|  |  | Do new things | Do things better | Stop doing things |
| Type of benefit valuation | Financial |  |  |  |
|  | Measurable |  |  |  |
|  | Observable |  |  |  |

*Table 1 Adapted benefits grid by Nielsen and Persson [11]*

The BC model may be defined through five steps: 1) Define drivers and investment objectives, 2) identify benefits, their measures and owners, 3) structure the identified benefits, 4) identification of cost and risks, 5) acceptance [11]–[13].

The BC model's primary goal is to nuance an organisation's perspective on the benefits of IT investments while also considering risks. In addition, it establishes a multidimensional model of thinking and structurisation of benefits concerning a specific IT investment, with a supplementary focus on reducing documentation and simplifying it [13, p. 4]. The BC model furthermore works as an agreement, which should handle the realisation

of benefits. However, we will not utilise the BC model as an agreement or checklist for realising benefits but as an approach to challenging the dissemination and structurisation of benefits when considering AT as an IT investment. Therefore, our utilisation of the BC model in this article is not to decide whether AT is a viable investment for EV, but as a tool for challenging their mode of thinking when considering benefits.

# 3 Methodology

This section will present our methodological considerations, which will have their foundation in the AR method [14]. AR focuses on practical problems in the real world. Solving the identified problems should bring new knowledge to the existing literature during an iterative process, visiting and revising the existing literature on AT and value creation and the actual problem in the organisation. While utilising AR to create new knowledge, we started by seeking knowledge in the existing literature concerning AT and value creation and participating in collaboration with EV for seven months. EV approached us as one of us was participating in an internship at EV and had shown interest in AT and theories of value creation and asked if we wanted to collaborate in disseminating AT as an investment for EV's product owners and consultants.

Furthermore, we will in this section elaborate on how we have researched the benefits of AT in the context of EV, utilising the concept of AR to gather knowledge in the existing literature, and use the literature findings to discuss and research practitioners of AT. We start by elaborating on who EV is and then focus on the research activities we conducted during these seven months of collaborating with EV.

## 3.1 The Company – Elbek & Vejrup

EV is a software consultancy- and development organisation that makes customer-adapted Microsoft dynamics 365 Business Central solutions. EV has offices in Aalborg, Herning, Kolding, and Copenhagen. However, their main office is in Aarhus, where we engaged with their consultants and POs. EV was founded in 1996 and had been expanding rapidly, averaging a 10% growth in employees per year. The expertise of different employees and specialists is used on loan to the departments needing specific knowledge or expertise. The increase in growth and complex solutions has made EV more determined to expand AT to all teams, potentially increasing the testing efficiency. Due to EV working with different industries, all require different solutions. Testing and ensuring product quality is crucial but very resource-heavy – a focal point for EV to mitigate by expanding their use of AT.

The management of EV decided before giving orders to the various departments that they would provide a forum of knowledge-sharing for all members of EV to become aware of the vision regarding the implementation of AT. We had meetings with the CEO, head of HR, and the leading consultant regarding knowledge-sharing and internal training. We were tasked with providing barely enough insight into AT concerning information systems and articulating it to the POs and consultants of EV. We were told to reflect and connect AT with how they are developed in MS365BC in relation to [15], as Vugt would educate and train the developers

in the coming months. This came to be the foundation of further implementation of AT at EV, or at least debates concerning AT. During this research activity, we gathered information on the overall knowledge regarding AT at EV, providing insights into how we were to disseminate the topic for the scheduled presentation and panel debate. We had to disseminate the benefits of AT from an idealistic and organisational perspective, rather than on time and cost to realise the benefits.

Due to EV's approach to knowledge sharing among employees, they facilitated for us to conduct a presentation with the following panel debate, giving us the chance to disseminate the benefits of AT. This was facilitated for their *old school consultants,* who lacked knowledge of or were strictly sceptical of AT, and consultants from a new wave of consultants, who had already worked with AT and advocated its expansion.

## 3.2 Research activities

Our research activities may be deemed to have consisted of 1) understanding and identifying benefits by using the BC model, 2) presentation of unstructured benefits to EV and observing the focus of the panel debate among EV consultants and POs, 3) Structure benefits and adapting the used model on the premises of our findings, revising and adapting our dissemination for the subsequent intervention, 4) presentation and discussion of structured benefits with POs of EV, 5) analyse gathered data of the discussions concerning our research focus, and 6) present and disseminate our findings concerning our research question.

As a part of our utilisation of the BC model, we addressed the first step by identifying a definition for the area of concern: AT. Our research of AT literature pointed us toward Dustin et al.'s definition concerning *automation of software testing activities: development and execution of test scripts, verification of testing requirements and use of test tools* [2], [9]. This definition helped us clarify the drivers and objects of AT as an investment.

While working with this definition, we identified nine categories concerning distinct aspects, as presented in 2.1 Benefits of automated testing. Furthermore, we identified the measures and owners of the benefits, product owners (PO) and management, while we divided the measures into the categories; time, money, and quality. We considered time and money as a resource perspective for the financial and measurable benefits, which may influence the quality perspective as an observable benefit.

**Our first intervention** concerned a presentation where we presented different aspects of AT, where the focus was on *what, why, how and who*, emphasising *why*. EV had booked all consultants and POs to attend this presentation, and approximately 50-60 employees attended on location and 5-10 attended online via MS Teams. Our presentation of *why* was based on the identified benefits of AT. We also had prepared questions for the attendees to discuss for the second part of our intervention; the panel debate. We conducted the presentation and asked our first question, challenging the attendees on their knowledge and experience with AT for them to disseminate the benefits of AT from a practitioner's perspective to their colleagues. During the panel

debate, we let the attendees discuss among themselves and only answered questions if they were directly addressed to us. When silence occurred, after debating a question, we would provide them with other questions, extending the debate for 1,5 hours. After the presentation and debate, we observed and gathered several insights and experiences, which we considered while revisiting the BC model and working on the next steps of the model: the structuring.

We structured our benefits by utilising the gathered empirical data concerning the research of Rafi et al. [7], Lindholm [8], and [16], who have gathered, compared and analysed benefits; focusing on how to directly measure the benefit, e.g. pace of test-time, and test-coverage. When structuring the benefits, we revisited the literature, which had identified the benefit, to explore and understand the contexts of the benefits for us to explore the mode of thinking. More concise: we considered each benefit concerning which new benefits they may bring about, which things the benefit might stop, and which things the benefit might do better, and generated the value map, Table 2, during this process. The measures and owners concerning each benefit were further deduced from the utilised and mentioned literature for us to strengthen the structuration of benefits.

We consider Table 2 to be a proposition of a value map regarding the benefits of AT. Our proposed value map is comprised of a modified benefits grid compared to its original form[11]–[13], where the focus has been on "which new things, or enhancement of present processes that implementing AT may bring about, and which things the organisation may stop doing". Furthermore, the grid is divided into different types of benefits, the financial, which concerns financial and quantifiable benefits; the measurable, which is comprised of evaluable benefits; and the observable measurements, which consider experience and judgement concerning predetermined criteria for realisation of benefits[13]. We provided identifiers for each benefit in Table 2, naming them F, M or O depending on their scope, and a number to distinguish the benefits.

In our process of using the BC model for exploring AT from an organisational perspective, we did not work with the actual cost of implementation but merely an indexed approximation. The cost of implementing AT is essential, as it may be alpha or omega concerning deeming an investment feasible. Identification of cost is a part of the fourth step of the BC model, where an absolute price is sought after, but in the case of AT, we deem it difficult to be precise and deem a relative price suitable. We base this on various dialogues with team leaders and the EV management (After this intervention, multiple managers and POs approached us to discuss AT's value). We found out that they charge/spend approximately 15-25% more to their customers when AT is built along with their personalised MS365BC solution. We believe that it is vital to address that this extra cost only occurs if there are no AT or test cases which may be reused from earlier projects, and everything must be developed from scratch. Van Vugt indirectly supports this finding, as he states that an MS365BC project, on average, exceeds its budget by 25% due to bug fixing after the solution is live, and the bugs may be found earlier if AT is implemented [15, pp. 6–7]. The other part of step four in the adapted BC model focuses on identifying risks, which we have combined with AT limitations. We believe that identifying drawing on Van

Vugt's risk perspective could use a contextual elaboration. Van Vugt identifies "*Why you should not implement AT*" as 1) "Operational excellence is mitigated due to high cost, 2) we are not used to working this way, 3) customers usually test the functionality, 4) lack of staff, 5) too busy with other tasks, 6) too big of an MS365BC solution to know where to start" [15, pp. 5–6].

The fifth part of the adapted BC model calls for identifying key actors. In this step, we decided to focus on the management and POs; the management decides the size of the investment and which direction and vision to follow, and the PO decides how the developer's time should be used to ensure the developed product. Concerning the implementation of AT, the management needs to start the process by revising the direction and vision and allocate resources to the PO to decide on the change of direction and vision and begin training for the team. We believe that the PO needs to be involved in the decision-making process, as the PO knows the range of the system/product and the team's skill level. Therefore, we chose to invite three POs and managers of EV to explore the benefits of AT and the articulation and mode of thinking of AT benefits.

On this foundation, we utilised the categorisation of financial, measurable, and observable benefits to be more precise about the identified benefits when disseminating them to the POs we had invited for interviews.

**We conducted a second intervention**, where we interviewed 3 POs concerning our structured benefits and discussed the benefits of AT about their technical thoughts as well as their practical impact on the organisation. The interviews were scheduled to last 1,5 hours. We had made printouts for each interview, providing a common language for how we articulated the identified and structured benefits. The interview was initiated by briefly explaining the value map and how we would like them to consider the benefits they would be presented with during the interview. We furthermore explained the categories *of Financial benefits, Measurable benefits, and Observable benefits* about whether the benefit provides *new things, better thing, or stop things from happening* for them to get into the mindset of consciously using the BC model to articulate and disseminate benefits of AT. During the interviews, we allowed the POs to reflect upon the presented and structured benefits concerning their experience with AT. We asked them whether they agreed, disagreed or were indecisive toward each structured benefit. If the POs did not reflect upon their answer or gave an elaborative answer, we would prompt them to elaborate on their reasoning for their answer. We prompted them as our intervention sought to sow the seed of positively manipulating their mode of thinking regarding benefits to gather insights into how a structured- and multidimensional approach to benefits might influence the organisation's understanding of value.

We believe these activities would provide insight into how organisations may change their mode of thinking regarding benefits. Furthermore, we initiated an analytical process of considering our findings of the second intervention to understand how our AR approach may have challenged the POs' dissemination of benefits. We believe this approach may bring about an attempt to challenge the existing literature on AT benefits concerning how benefits are articulated and processed when considering practitioners' approach to structure vs no

structure. The utilisation of the BC model may bring about a more nuanced approach to the existing literature on AT benefits while simultaneously addressing the BC model's approach to handling IT investments; not focusing on the feasibility of the investment or the documentation that the BC model may be viewed as, but the applicability regarding AT.

# 4 Findings

In this section, we will elaborate on the findings, addressing the problem at EV and our two interventions.

## 4.1 The problem situation at Elbek & Vejrup

Our meetings with the CEO and Head of HR revealed the consensus of AT at EV. Some consultants did not care for AT and did not want to work with it, as it would only slow them down in their implementation of new solutions to customers while having to increase the cost of the project. Various departments have already been using AT for years, whereas others have not. AT was often articulated as a time and money spender, leaving no room for developing new features, as many of the MS365BC solutions are 15+ years old, full of complexity, related and interrelated functionality. On the other hand, the developers, especially the younger generation, liked the idea of working with AT, and those who already worked with it advocated it to the other developers. The departments, which had similar customers, whose structure in the solutions were more or less identical, and already was working with AT were more inclined to accept the management's vision at EV, thereby accepting the enhancement of resources allocated to and expansion of AT at EV.

## 4.2 First intervention – unstructured dissemination

After a month of researching AT and relating it to MS365BC solutions, we had our first intervention during the event "EV-directions", a knowledge-sharing day, where all the consultants and POs were invited to listen to our presentation and the panel debate. The intervention started with us giving a presentation concerning What is AT? Why AT? How do we implement AT? Who implements AT? and benefits (unstructured) concerning AT. The following debate mainly concerned the implementation process of AT regarding functionality and scope. Shortly after the panel debate started, we observed three groups forming among the consultants and POs; a group focusing on the benefits, another on limitations, and the third which had no strict opinion but wanted to know more. We further observed that the debate about the product's quality with AT's implementation was quickly acknowledged. The rest of the debate concerned three questions: "*How much does it cost – not relative price, but absolute*", "*How much time does it take – not relative time, but absolute*" and "*How much of the system should be covered by AT?*".

This intervention brought about several indications of a potential difference in the research we had studied, and the benefits we had identified; 1) product quality should not be contested regarding AT vs MT (manual testing), 2) risks of AT was not mentioned during the debate, but the focus revolved around the limitations, 3) "the customers pay", so we need to be clear and concise about economic- and time-reducing benefits and 4)

how much of the MS365BC solutions should be covered by AT. When we revisited our unstructured benefits, these deductions became the four focal points for us. We, therefore, decided that we needed to revise and adapt our utilisation of the BC model and include limitations for us to understand "how, when, and why to implement AT in MS365BC solutions. We revised our benefits and planned meetings with POs to challenge the POs regarding a structurisation of the benefits, considering each benefit through the matrix of the benefit grid. While considering the benefits after conducting the first intervention, we managed to structure the benefits we observed in AT literature, as presented in Table 2. All 13-benefits concern money, time, coverage, and quality, as these overall categories were observed in our literature study and during the first intervention.

| | Do new things | Do things better | Stop doing things |
|---|---|---|---|
| **Financial benefits** | **Benefit F1:** Cost may be defined at the beginning of projects [9], [15], [17]–[21]<br>**Measure:** Resources allocated during the lifecycle | **Benefit F2**: Resource allocation for testing is reduced [9], [15], [17], [21]<br>**Measure**: Cost of error correction<br>**Owner**: Management | |
| **Measurable benefits** | **Benefit M1**: Acceptance of code before merging new functionality into the system [8], [15], [18], [22]<br>**Measure**: Time spent at the beginning vs during lifecycle<br>**Owner**: Management and product owner<br><br>**Benefit M2:** Testing becomes 100% consistent where there is no human testing [9], [15], [18], [19], [21], [23], [24]<br>**Measure**: Deviations in test<br>**Owner**: Product owner<br><br>**Benefit M3**: Tests and test cases may be reused for other tests and becomes easier to develop [9], [15], [19], [21], [23], [24]<br>**Measure**: Time spent on developing new AT<br>**Owner**: Management and product owner | **Benefit M4**: The system will contain fewer errors during its lifecycle [9], [19], [21], [25], [26]<br>**Measure**: Number of errors reported, and time spent fixing errors<br>**Owner**: Management and PO<br><br>**Benefit M5**: Testing the system becomes less prone to human errors [7]–[9], [19], [22], [27]<br>**Measure**: The same errors will be found every time the same test is executed<br>**Owner**: Management and PO<br><br>**Benefit M6**: Less time is used on testing, which frees up time for other tasks [15], [19], [21]–[24]<br>**Measure**: Time allocated to testing, and time spent on non-testing tasks<br>**Owner**: Management and PO<br><br>**Benefit M7**: The testing of the system becomes more reliable [7]–[9], [19], [21]<br>**Measure**: Number of deviations recorded<br>**Owner**: PO | **Benefit M8**: Deviations in the same executed tests stop [7]–[9], [19], [21]<br>**Measure**: Deviations recorded during tests<br>**Owner**: PO |
| **Observable benefits** | **Benefit O1**: Quality is maintained when implementing new functionality [7], [15], [19], [27] | **Benefit O2**: Improved quality in the system and the specific functionality of the system due to more undersized and inexperienced human/manual testing [7], [19], [21] | **Benefit O3**: There will no longer be deviations when the |

| | | | same test I executed [7], [9], [21] |
|---|---|---|---|
| | **Measure**: Number of errors observed while implementing new functionality<br>**Owner**: Management and PO | **Measure**: Numbers of errors found during testing<br>**Owner**: Management and PO | **Measure**: Number of errors recorded concerning specific tests and functions<br>**Owner**: PO |

*Table 2 Value map of AT benefits*

## 4.3 Second intervention – Structured dissemination

Our second intervention consisted of dialogues with POs from EV where two of the POs (PO2 and PO3) were product owners/managers, and one was strictly PO (PO1). These informants were chosen because of their knowledge and experience with AT and their position at EV; we believe it was important to have fewer participants with deep knowledge rather than many people with lesser AT knowledge as to how our approach to AR iterates between academia and organisations. Furthermore, due to our interview approach of conducting semi-structured interviews, we had a very different perspective on almost every benefit, indicating how the BC models approach to structurisation may influence the mode of thinking in organisations. The main findings of the second intervention will be presented, summing up the POs' experience and dissemination of structured benefits in Table 3, which presents the financial benefits, Table 4, which elaborates on measurable benefits, and Table 5, which considers observable benefits.

| ID/ respondent | PO1 | PO2 | PO3 |
|---|---|---|---|
| F1 | Agree, but only the core functionality | Disagree; too many dissimilarities | Agree; if the system is static and the staff can write the correct test |
| F2 | Agree; it depends on how much time we spend and the value of the tests | Agree, if the cycle is not too short | Agree, but the proper test needs to be written, and the lifecycle is not always known. |

*Table 3 Financial benefits according to the POs*

When asked about the financial benefits, F1 and F2 in Table 2, PO1 focused on how his team had been able to define the cost of their core- and business-critical functionality, especially regarding their system, which handles consumption bills before initiating new and updated releases. PO1 had also experienced that the dissimilarity of their customers' solutions made it difficult, if not impossible, to define the cost of non-core and non-reused functionality. PO2 also focused on the dissimilarities of the solutions, disagreeing with benefit F1on the premise that each customer is different by nature, thereby making it impossible to define the cost, thus rejecting the premise that AT establishes a new financial benefit. Finally, PO3 pointed out that it had been possible for their team, but only with static functionality of their system that has been tested extensively, thereby making his team understand the functionality fully; if these premises are not present, then PO3 rejects F1.

When we presented benefit F2 to the POs, they all agreed that the allocation of resources was reduced and thereby better after implementing AT. PO1 focused on freeing up time for his team during the lifecycle and thereby finding time to develop other tasks ordered by the customers. PO2 focused on the length of the lifecycle and believed that the length was alfa or omega for comparing "how much better" the AT is when working with MS365BC. Finally, PO3 focused on static functionality and whether the team has built the correct test; only if these premises are met then PO3 agree that AT are reducing the resource allocation for testing. When going through our transcripts of the interviews, we noticed that not only did the POs focus on the category of financial benefits, but they also included "new, better, or stop" in their answers. However, PO1 did not consider the financial aspects of resource allocation in the same manner that PO2 and 3 did but were much more focused on categorising "new, better, or stop".

|    | PO1 | PO2 | PO3 |
|----|-----|-----|-----|
| M1 | Indecisive, as manual inputs are required | Indecisive, due to considerations regarding strategy | Agrees, change requires manual input |
| M2 | Agree | Indecisive, leaning towards agreeing | Agree, both new and better. No time save short term |
| M3 | Agree; wrong datatypes usage can be bad | Long-term: agree Short term: disagree | Agree; experience is needed to adapt changes to tests |
| M4 | Agree | Agree, but mentions the limitation of inexperience. Change requires evaluation | Agree, but believes the tests need monitoring |
| M5 | Agree | Disagree; it needs monitoring as it is hard to cover all functions with at, and more experience is needed. | Indecisive; errors will happen, but fewer will occur. |
| M6 | Agree: "*most important benefit for us*." | Agree | Agree; hard to predict work tasks |
| M7 | Agree; it reduces manual testing but does not stop it | Agree, as manual testing may create issues, if we are not precise and stringent | Agree, when we have the suitable test |
| M8 | Agree; it requires experience before it becomes efficient | Agree; the more humans involved; the more deviations will happen | Disagree |

*Table 4 Measurable benefits according to the POs*

When we made the POs consider the measurable benefits and articulate the benefits we presented, we gathered what may be deemed more nuanced answers and dissemination. For example, when asked about M1, PO1 was indecisive, as they had experienced that their solutions used a lot of temporary tables, making their complex to the extent that they do not have enough experience with AT to comprehend the functionality. PO1 further focused on that AT indeed provide the aspect of "new", as the functionality their team may comprehend regarding dependencies of temporary data allows for an acceptance test, thereby deciding if the functionality is to be implemented without further testing. PO2 also focused on dependencies of temporary data and proposed developing a model to handle "when and what to automate" concerning complexity and time spent on the test. When the test is within the model's threshold, PO2 agrees with benefit M1. PO3 was not as focused on the

temporary data and agreed with benefit M1, emphasising that this benefit is new and that the benefit might overflow to other benefits in the "making things better" category.

When we talked about benefit M2, there was a slight consensus that the fewer people involved in testing, the more consistent the testing will be, and all three POs agreed that the benefit was a new aspect of testing. However, PO2 argued and had experienced that the removal of humans in testing may create false positives, and maybe the test will not find the errors, making PO2 indecisive about whether benefit M2 is a benefit. On the other hand, PO3 was more focused on the time-saving aspect of removing human testing and believed that this benefit might be structured as both a "new and better" thing. M3 caused consensus with all three POs, where experience and maturity of AT were the focus, while all arguing that this benefit may only be characterised as a "new thing" and over time, this might be the most beneficial "new thing".

When asked about M4, PO1 pointed out that for their team, this did not only bring about an enhanced process but also a new process. PO1 elaborated that their developers did not test their code extensively before implementing AT. However, they only developed what the consultants had asked for and let the consultant determine if the code could be accepted. Now they have "a button" for testing if their new tests can be accepted. PO2 argued that his team needed more experience to catch everything, making this benefit and the concept of "better things" a question of zero-sum. However, PO2 acknowledged that the process has improved in core functionality, where they have a comprehensive AT. PO3 agreed with the enhancement of testing but mentioned that MT would not stop.

None of the POs' answers corresponded with the others regarding benefit M5, where PO1 had experienced a much better process. PO2 again considered the benefit of its category as a zero-sum equation, deeming "if not everything is made better, I do not agree with the benefit". On the other hand, PO3 considered this benefit a positive-sum benefit, where if it made something better, he agreed; but believed that the AT needed time before the "better" would be fully realised.

When we presented M6 as a benefit to improve the process of time-spent, all three POs did agree, emphasising that the AT may need maintenance sometimes. However, when measured over time, none would challenge the time they spent on testing now versus before they implemented AT.

When we asked the respondents to consider benefit M7, PO1 focused on reliability as a zero-sum, where AT, in this regard, will not stop anything or do anything new but simply enhance a process. The approach of PO1 was backed by PO2 and 3, as all of them articulated that AT will not bring about a new measurable benefit or stop an unbeneficial "thing", as deviations may happen. However, there may also be tests where no deviations may be measured. All POs articulated this benefit when asked about benefit M8, where PO1 argued that they only find deviations in tests when humans are involved. PO2 considered variables and changes; both agreed with the structuration of the benefit and that deviations overall stop if specific conditions are met. Finally, PO3

argued that their solutions would have deviations, as their solutions are very dynamic, and they change their AT often to comprehend how their customers want their solutions updated.

| | PO1 | PO2 | PO3 |
|---|---|---|---|
| O1 | Agree, but only core functionality and critical functions, as it is static | Agree; fewer errors equal better quality | Agree if the correct test is made |
| O2 | Agree; more experience can increase the quality | Indecisive | Agree |
| O3 | Agree, but on static and core functionality | Indecisive, not beyond core functionality | Agree, but a long-term goal |

*Table 5 Observable benefits according to the POs*

When asked about O1, all POs agreed that AT could enhance product quality. However, they remained critical towards realisation hereof. PO1 stressed that improvements in product quality would be on core and critical functionality. PO1 further explained that they had not received customer complaints since implementing AT on core functionality.PO2 agreed with the concept of O2 and discussed the incentive to expand AT beyond core functionality. Moreover, this would make an expansion of AT possible. PO3 added that any improvements to quality would only occur if they had the suitable parameters for the test and that if the parameters are met, AT secures quality.

When asked about the benefit of O2, PO1 talked about how they would receive complaints about blocks in their daily flows. However, after implementing AT, the complaints have decreased and added that more time using AT on solutions might further solve the issue. PO2 was indecisive regarding O2 but talked about how it made sense to expand AT beyond core functionality to ensure customers do not experience errors. The universal solutions make AT possible. PO3 stated that proper evaluation of what they should implement AT on and evaluating if the scripts are in sync with customer usage. Moreover, this only comes with experience with implementing AT.

When asked about O3 deviations in test execution, the POs all agreed that deviations would no longer occur. PO1 talked about how this stop of deviations created some new safety in their product, but the stop was strictly for the core functionality. PO2 called it a theoretical yes but regarded the complexity of the systems as incomprehensible. PO2 further elaborated that it would not be possible to cover all but agreed with PO1 that on the core functionality, deviations stopped PO2 reasoned that the cause of deviations in the periphery happens because the test is not executed the same way every time. Finally, PO3 mentioned the benefit of O3, a long-term goal for them, and talked further about how to archive stop deviations by becoming more experienced, making test scripts, and gathering good information from their customers. Moreover, when they reached a good flow in making the test scripts, deviations in the test would no longer occur.

We believe it is essential to address our overall findings of this second intervention. We have observed several differences after we presented a framework for disseminating and structuring benefits at EV. We observed that

the more we discussed the structured benefits of table 2, using the terminology of the BC model with the respondents, the more they used the multidimensional benefits grid when discussing a potential benefit. Furthermore, we observed that when asked about a benefit from Table 2 Value map of AT benefits, or asked to elaborate on their answer, the POs related the benefits to limitations or conditions, addressing the identified benefits in the existing literature, but disseminating them in a structured and multidimensional approach.

All three POs used the BC models structurisation of types of valuation and change when addressing the observable benefits and stated that using the framework for discussing benefits had helped them clarify and disseminate their experiences of AT. Two of the three POs asked if they could keep the printout of table 2, as they believed they could use this mode of structuring benefits for other purposes than AT.

## 5 Discussion

Understanding the benefits of Automated Testing has been studied since the 1990s [2], [22], and there is consensus concerning which benefits AT may bring to software development [1], [7]–[9]. The nine categories highlighted by Rafi et al. [7], along with the empirical observations of Karhu et al. [9], have been the foundation for this article's purpose. The focus and purpose have been understanding and challenging the dissemination of Automated testing benefits in an organisational context through our research question: *"How can we impact the dissemination of Automated testing at a software company by using a business case model?"*.

We approached the research question by utilising Nielsen and Persson's adapted Business case model [11], [13] through an Action Research approach, attempting to impact the organisation without providing consultancy. The concept of disseminating the benefits of Automated Testing may be deemed to be a well-studied topic, where the existing literature has reduced the overall dissemination of Automated Testing benefits to nine categories, all focusing on time, money, and quality [1], [7]–[9], [28].

When we presented the nine benefits to EV in our first intervention, they did not focus on the benefits but on the resource expenditure and how they could articulate implementation or an expansion of Automated Testing in their systems and convince their customers to pay. When we conducted our second intervention, utilising the concept and approach of structuring benefits for multidimensional dissemination, we observed that the dissemination of the benefits no longer was based on *"how much it costs and how do we present it to our customers?"*. Instead, we observed that the focus of the product owners was "type of valuation and type of change" concerning each of the same benefits, which we had presented five months earlier, expanding the scope of the discussion of the benefits. We further observed that the second intervention enhanced the utilisation of limitations and conditions, equal to the presented limitations of Rafi et al. [7] and the conditions observed by Karhu et al. [9]. As a result, the scope of each benefit was articulated much clearer and helped the respondents relate the benefit to potential pitfalls. Thereby, we may conclude that utilising the Business case model impacted EV's product owners' dissemination and articulation of benefits regarding Automated Testing

towards a multidimensional and structured approach. Furthermore, the utilisation has brought about an understanding of using limitations and conditions equal to those existing literature on the benefits of Automated Testing mentioned when reflecting on the presented value map, Table 2.

When utilising the Business case model in our Action Research approach, we did not consider risks, as they were not mentioned during our first intervention. We, therefore, adapted our approach to consider limitations and conditions instead of risks. We do not conclude that the Business case model needs to implement this adaptation, as it would require more research.

However, we argue that an essential contribution of our Action Research is that limitations and conditions should be considered when analysing Automated Testing through the adapted Business case model. Furthermore, using limitations and conditions may bring about reflection when considering potential benefits, as we observed that presenting benefits in a structured approach brought about reflections of limitations and conditions. We believe this statement needs further research; we, therefore, propose workshops should be conducted. The workshops should revolve around product owners and activities of identifying and structuring benefits of Automated Testing using the Business case model's adapted benefits grid while considering limitations and conditions to conduct empirical observations.

# Bibliography

[1]  D. L. Asfaw, "Benefits of Automated Testing Over Manual Testing," *Int. J. Innov. Res. Inf. Secur.*, vol. 2, no. 1, p. 10, 2015.

[2]  E. Dustin, J. Rashka, and J. Paul, *Automated software testing: introduction, management, and performance*. USA: Addison-Wesley Longman Publishing Co., Inc., 1999.

[3]  E. Dustin, T. Garrett, and B. Gauf, *Implementing automated software testing: how to save time and lower costs while raising quality*. Upper Saddle River, NJ: Addison-Wesley, 2009.

[4]  B. Pettichord, "Seven steps to test automation success," 1999.

[5]  S. Berner, R. Weber, and R. K. Keller, "Observations and lessons learned from automated testing," 2005, pp. 571–579. doi: 10.1109/ICSE.2005.1553603.

[6]  K. C. Laudon, J. P. Laudon, and J. P. Laudon, *Management information systems: managing the digital firm*, Fifteenth edition, Global edition. Harlow London New York NY: Pearson Education Limited, 2018.

[7]  D. M. Rafi, K. R. K. Moses, K. Petersen, and M. V. Mäntylä, "Benefits and limitations of automated software testing: systematic literature review and practitioner survey," in *Proceedings of the 7th International Workshop on Automation of Software Test*, Zurich, Switzerland, Jun. 2012, pp. 36–42.

[8]  D. Lindholm, "Economics of Test Automation : Test case selection for automation," *undefined*, 2019, Accessed: Mar. 10, 2022. [Online]. Available: https://www.semanticscholar.org/paper/Economics-of-Test-Automation-%3A-Test-case-selection-Lindholm/c217a2ca1f9f0f40f3560660c3a97aa6c6ea9d1b

[9]  K. Karhu, T. Repo, O. Taipale, and K. Smolander, "Empirical Observations on Software Testing Automation," 2009, pp. 201–209. doi: 10.1109/ICST.2009.16.

[10] M. Cui and C. Wang, "Cost-Benefit Evaluation Model for Automated Testing Based on Test Case Prioritization," *J. Softw. Eng.*, vol. 9, no. 4, pp. 808–817, 2015, doi: 10.3923/jse.2015.808.817.

[11] P. A. Nielsen and J. S. Persson, "Useful business cases: value creation in IS projects," *Eur. J. Inf. Syst.*, vol. 26, no. 1, pp. 66–83, Jan. 2017, doi: 10.1057/s41303-016-0026-x.

[12] J. Ward, E. Daniel, and J. Peppard, "Building Better Business Cases for IT Investments," *Mis Q. Exec.*, Jan. 2008, Accessed: Dec. 14, 2021. [Online]. Available: https://www.academia.edu/17870298/Building_Better_Business_Cases_for_IT_Investments

[13] P. A. Nielsen and J. S. Persson, *Business-case metode for kommunale IT-investeringer, version 2.0*. Software Innovation Publisher, 2013.

[14] J. McKay and P. Marshall, "The dual imperatives of action research," *Inf. Technol. People*, vol. 14, no. 1, pp. 46–59, Mar. 2001, doi: 10.1108/09593840110384771.

[15] L. van Vugt, *Automated testing in Microsoft Dynamics 365 Business Central: efficiently automate test cases for faster development cycles with less time needed for manual testing*. 2021. Accessed: Feb. 16, 2022. [Online]. Available: http://public.eblib.com/choice/PublicFullRecord.aspx?p=6826438

[16] G. V. Murazvu, "Software Testing: An Analysis of the Impacts of Test Automation on Software's Cost, Quality and Time," 2020.

[17] M. S. G. Devasena, G. Govindasamy, and M. Valarmathi, "Automated and Optimized Software Test Suite Generation Technique for Structural Testing," *Int. J. Softw. Eng. Knowl. Eng.*, vol. 26, pp. 1–13, Feb. 2016, doi: 10.1142/S0218194016500017.

[18] L. du Bousquet and N. Zuanon, "An Overview of Lutess: A Specification-based Tool for Testing Synchronous Software.," Jan. 1999, pp. 208–215. doi: 10.1109/ASE.1999.802255.

[19] B. Haugset and G. Hanssen, "Automated Acceptance Testing: A Literature Review and an Industrial Case Study," Sep. 2008, pp. 27–38. doi: 10.1109/Agile.2008.82.

[20] L. Shan and H. Zhu, "Generating Structurally Complex Test Cases By Data Mutation: A Case Study Of Testing An Automated Modelling Tool," *Comput. J.*, vol. 52, Aug. 2009, doi: 10.1093/comjnl/bxm043.

[21] J. Kasurinen, O. Taipale, and K. Smolander, "Software Test Automation in Practice: Empirical Observations," *Adv. Softw. Eng.*, vol. 2010, Jan. 2010, doi: 10.1155/2010/620836.

[22] D. Hoffman, "Cost Benefits Analysis of Test Automation," p. 13, 1999.

[23] J. A. Dallal, "Automation of object-oriented framework application testing," in *2009 5th IEEE GCC Conference Exhibition*, Mar. 2009, pp. 1–5. doi: 10.1109/IEEEGCC.2009.5734312.

[24] D. Banerjee and K. Yu, "Robotic Arm-Based Face Recognition Software Test Automation," *IEEE Access*, vol. 6, pp. 37858–37868, 2018, doi: 10.1109/ACCESS.2018.2854754.

[25] S. Stresnjak and Z. Hocenski, "Usage of Robot Framework in Automation of Functional Test Regression," Oct. 2011.

[26] G. Tassey and Gregory, "The Economic Impacts of Inadequate Infrastructure for Software Testing," May 2002.

[27] F. Saglietti and F. Pinte, "Automated unit and integration testing for component-based software systems," in *Proceedings of the International Workshop on Security and Dependability for Resource Constrained Embedded Systems*, New York, NY, USA, Sep. 2010, pp. 1–6. doi: 10.1145/1868433.1868440.

[28] V. Garousi and M. V. Mäntylä, "When and what to automate in software testing? A multi-vocal literature review," *Inf. Softw. Technol.*, vol. 76, pp. 92–117, Aug. 2016, doi: 10.1016/j.infsof.2016.04.015.