Preconditions for the Use of a Checklist by Enterprise Architects to Improve the Quality of a Business Case

van den Berg, Martin; Cordewener, Lars; Lago, Patricia; Slot, Raymond; van Vliet, Hans

2019

DOI (link to publisher)
10.1109/CBI.2019.00020

document version
Publisher's PDF, also known as Version of record

Preconditions for the Use of a Checklist by Enterprise Architects to Improve the Quality of a Business Case

Martin van den Berg
Dept. of ICT
De Nederlandsche Bank
Amsterdam, The Netherlands
m.j.b.k.van.den.berg@dnb.nl

Raymond Slot
Strategy Alliance
Amersfoort, The Netherlands
raymond.slot@strategy-alliance.com

Lars Cordewener, Patricia Lago, Hans van Vliet
Dept. of Computer Science
VU University Amsterdam
Amsterdam, The Netherlands
lars.cordewener@gmail.com, p.lago@vu.nl,
hans@cs.vu.nl

Abstract— **Enterprise Architecture (EA) is a mechanism to support IT investment decisions. What is lacking, is a practical tool like a checklist that can be used by enterprise architects to support IT investment decisions, and more specifically, the business case. The business case is a common practice to justify IT investment decisions. The objective of this research is to investigate how the use of a checklist can support enterprise architects to improve the quality of a business case. To answer this question, we conducted design science and developed and evaluated a checklist that can be used by enterprise architects in the assessment of a business case. The evaluation of the checklist was conducted by means of an experiment. The results of the experiment with 32 architects demonstrate that using a checklist as such does not lead to higher quality improvement suggestions in the assessment of a business case as compared to not using a checklist. In other words, a checklist alone does not make the difference. Discussion of the results of the experiment with focus groups reveals several preconditions for the use of a checklist. In any case, the use of the checklist reduces the risk of overlooking valuable insights to improve the quality of the business case.**

Keywords- Enterprise architecture, checklist, business case, IT investment decision, design science

I. INTRODUCTION

The IT investment decision is the decision to approve or reject an IT investment. IT investment decisions "select and fund initiatives and address how much to spend, what to spend it on, and how to reconcile the concerns of different stakeholders" [1]. Not every IT investment is successful. The CHAOS report indicates that in 2015 only 29% of projects were successful, i.e., delivered on time, on budget, and with a satisfactory result. 19% of all projects failed and 52% were challenged [2]. The high failure rates of projects indicate that the IT investment decision is a difficult and risky decision. Different tools have been developed to mitigate the risks of an IT investment decision. The business case is one of them. Its purpose is to guide management in making the right IT investment decision.

Enterprise Architecture (EA) has similarities with the business case: it is a mechanism to support IT decision-making [3]. Different studies demonstrate that EA delivers insights [4, 5, 6, 7]. These insights are intended to be useful to different stakeholders, including decision-makers. EA insights should support decision-makers to make better decisions, including the decision to approve or reject a business case for an IT investment. In a previous study, we found evidence that EA can add to the quality of IT investment decisions [8]. In this earlier study we compared organizations with the highest quality of IT investment decision outcomes (top performers) with organizations with the lowest quality of IT investment decision outcomes (bottom performers). We found that EA offers top performers more key insights than bottom performers in the preparation of IT investment decisions, and in particular strategic insights. That brought us to the idea whether we can operationalize these insights in a practical tool that can support enterprise architects to improve the quality of a business case. The intention of this study is to evaluate the use of a checklist for enterprise architects by which they can assess the business case. This should ultimately improve the quality of the IT investment decision and lead to more successful projects. Since the role of enterprise architects in business cases has not been thoroughly researched, we position this research as exploratory. Our research question is: *How can the use of a checklist help enterprise architects to improve the quality of a business case?*

The main contribution of this study is a theory for design and action [9]. The value of EA is in the application of the knowledge and experience of enterprise architects, i.e., in providing valuable insights to improve the quality of a business case. This research aims to evaluate the use of a checklist that supports enterprise architects in the assessment of a business case. Our research reveals that a checklist alone does not make a difference unless the following preconditions are met: 1) the checklist should be properly introduced and communicated; 2) the checklist should be part of a procedure that should be followed strictly by architects; 3) the user of the checklist should be trained in how to use the checklist and the concepts behind the checkpoints; 4) the user should have knowledge of the domain of the business case; 5) the checklist should be extended with checkpoints on benefits, costs, and objectives.

The paper is structured as follows: In Section II we present related work. The research method is explained in

Section III. In Section IV we elaborate on the way we executed the research, and in Section V we present the results. The results are discussed in Section VI. In Section VII we discuss the threats to validity and limitations. Section VIII summarizes the conclusion.

II. Related Work

A. Enterprise Architecture

EA should guide decision-making [10, 11]. Various literature reviews indicate that EA leads to better decisions [12, 13]. One of the goals of EA is reduced risk of future IT investments [14]. There is broad consensus that EA guides and informs IT investment decisions [15, 16, 17]. There is even consensus that the enterprise architect plays a role in the development and evaluation of the business case [14, 18]. However, the use of EA differs across organizations and EA has different meanings for both practitioners and researchers. Lapalme introduced the three schools of EA [19]. He distinguishes between enterprise IT architecting, enterprise integrating, and enterprise ecological adaptation. Each school is grounded in its own belief system. Like in Lapalme's enterprise integrating school, we consider EA as the link between strategy and execution. In this research we applied Greefhorst and Proper's definition of EA: "EA is a discipline that is able to create overview and insights needed to translate strategy into execution, enabling senior management to take ownership of the key decisions on the design of the future enterprise" [20]. We consider decisions to approve or reject a business case as key decisions since business cases are at the basis of extending or changing the enterprise.

B. Business Case

The purpose of a business case is to guide management in making the right IT investment decision [21, 22]. Over the years the business case has become one of the best practices to support IT investment decisions. According to Ward et al. "developing a business case for IT investments is common practice. 96% of respondents reported that they were required to produce some form of business case when justifying IT investments" [22]. Business cases "assess multiple dimensions of IT projects, including the technology being deployed, its impact on targeted users and its implications for corporate strategies and business processes; business cases hence determine the success of IT investments" [23]. The business case is a standard part of widely used project management methodologies like Prince2 [24]. In this research a specific real-life business case was used that was assessed by architects.

C. Checklist

A checklist has several advantages. It is a way to deal with the ever increasing complexity in a practical way and it helps to make knowledge practical and disseminate it [25]. Furthermore, a checklist makes decision-making less dependent on individuals. According to Kahneman "organizations are better than individuals when it comes to avoiding errors, because they naturally think more slowly and have the power to impose orderly procedures" [26]. The use of a checklist is a form of an orderly procedure. A checklist can make the decision-making process more rational. Dean and Sharfman found that rational processes were associated with effective outcomes; political processes were not [27].

According to Gawande a checklist must be 1) precise, not vague; 2) practical; 3) contain reminders of the most important and critical steps; 4) simple but exact; 5) contain the "killer-items": the steps which pose dangerous risks if they are passed over; and 6) limited: must fit on one page [25]. In this study, the goal of the checklist is not to avoid danger, but to mitigate risks of IT investments. According to Gawande there are different types of checklists: "do-confirm" checklists and "read-do" checklists [25]. A do-confirm checklist is where you confirm you have carried out the specified action. In case of a read-do checklist you read the item and then go do what's specified.

In order to gain the necessary perspective on what is considered a good checklist we performed a systematic literature review (SLR). By searching the literature for papers that mention the quality of a checklist, we extracted data on several key factors that contribute to the success of a checklist. The literature was dominated by publications describing a checklist used in surgery, however many of them described its implementation in a particular domain of surgery to provide variety. Interesting data surrounding the corpus of papers includes the increasing trend of publications in recent years, with a large increase in publications occurring since 2012. In addition, the corpus confirms Gawande's preference for do-confirm checklists, as 83 out of the total 103 papers used this type of checklist. In addition, the number of checkpoints used in the checklists exceeds the 9 points that Gawande suggests, with the corpus averaging around a total of 19 checkpoints. The review checked the criteria that each publication used for measuring the success of their respective checklist, which showed a general trend in behaviour and implementation, both of which were keywords assigned to the effects surrounding the general behaviour of end-users towards the checklist as well as the practice of introducing the checklist to the already existing task it is intended to improve. Lastly, looking at the actual results as presented in the publications, making use of the keywords, we saw that the positive results associated with introducing a checklist were mainly focused on behaviour, implementation, and communication. This shows that the checklist was not only effective in reaching their intended goal, or improving the process of reaching said goal, but also to improve the behaviour of those that would be using the checklist and their communication on the tasks they are performing.

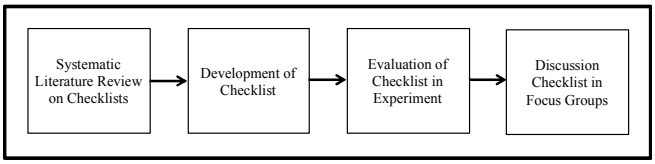

Figure 1. Research approach

III. RESEARCH METHOD

A. Research Approach

To address our research question, we apply design science [28]. Design science creates and evaluates artifacts intended to solve identified organizational problems. Its two main processes are "build" and "evaluate". The artifacts are constructs, models, methods, and instantiations. In our research, we intend to *build* a checklist for enterprise architects by which they can assess a business case. The checklist is based on earlier research identifying the insights that EA provides in the preparation of IT investment decisions [8]. The main focus of this research is on the empirical *evaluation* of the checklist. This evaluation took place in an experiment and the results were discussed in focus groups.

Figure 1 shows the overall research model. First, we studied literature on checklists. Second, we developed a checklist and third, we evaluated the checklist using an experiment with five groups. Fourth, we discussed the use of the checklist in focus groups. We elaborate on the second, third, and fourth step in the next subsections. The SLR was already discussed in section II.

B. Checklist Design

As part of this research a checklist was created with which architects can assess a business case. This checklist is intended as an instrument to be used by enterprise architects in an area where EA and IT investments come together; it can thus function as a boundary object between the communities of EA, project management, and senior management [29]. The design of the checklist was largely inspired by Gawande [25]. We applied Gawande's six criteria for a good checklist and created a do-confirm checklist. The first version of the checklist was based on earlier research and contains the key insights that EA can provide related to IT investment decisions [8]. The checklist is modelled like in Kahneman et al. and contains checkpoints where each one has some underlying questions to clarify the checkpoint and encourage the checklist user to look for answers to these questions in the business case [30].

C. Experiment Design

The model in Figure 2 shows the main variables relevant for the experiment, i.e., the checklist as the independent variable and the quality of the improvement suggestions for a business case as the dependent variable. The decision-making context and the EA context are considered as confounding variables.

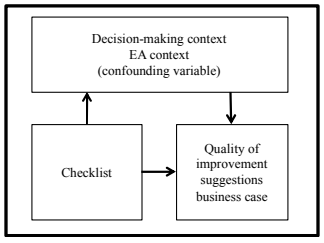

Figure 2. Research design of experiment

We followed Wohlin's et al. principles of experiment design [31]. According to Wohlin an experiment is "an empirical enquiry that manipulates one factor or variable of the studied setting. The starting point of such a design is that we believe that there is a relationship between a cause construct and an effect construct". The cause construct in this study is the use of a checklist, the effect construct is the quality of improvement suggestions for a business case. We have scoped this part of our research by means of a goal definition framework as:

Analyze the improvement suggestions for a business case of an IT investment, *for the purpose of* evaluation of the added value of a checklist, *with respect to* the quality of improvement suggestions for the business case, *from the point of view* of the decision-maker, *in the context of* architects assessing the business case.

This study is a so-called "multi-test within object study", which means that we study a single object, i.e., one business case, across multiple subjects, i.e., architects. Treatment one is architects using a checklist, and treatment two is architects not using a checklist in the assessment of a business case. This leads to the following hypotheses:

H0: there is no or a negative difference in the quality of the improvement suggestions between treatment one and two.

H0: $\mu\ S_{checklist} \leq \mu\ S_{withoutchecklist}$

Where μ is the average, S is the score of the improvement suggestions of the business case, $S_{checklist}$ is the score of the assessment with checklist and $S_{withoutchecklist}$ is the score without checklist.

H1: there is a positive difference in the quality of the improvement suggestions between treatment one and two.

H1: $\mu\ S_{checklist} > \mu\ S_{withoutchecklist}$

We analyze the results and statistically approve or reject the hypotheses by means of a t-test.

The object of this experiment is a real-life business case from the Dutch Government [32]. This business case was chosen because it was publicly available and because there was plenty room to seek improvement. The business case contains 42 pages including nine appendices. The requested investment is EUR 1.8 Million. The aim of the investment is to improve the System Catalogue of the Dutch Key Registers in terms of usability and content.

The subjects involved in our design are architects. We approached different networks of architects and consultancy and IT organizations that employ architects. Five of these networks and organizations collaborated. In the end 33 people participated across 5 groups. The participants had to assess the object and develop improvement suggestions.

The quality of the improvement suggestions of the participants was determined by a panel of five experts who were supposed to act as decision-maker. They were asked to

rank all improvement suggestions with a score between 1 and 5, where 1 is the lowest and 5 the highest score.

The decision-making context and the EA context are possible confounding variables that may influence the results of the experiment. Our experiment was conducted in an off-line situation, under controlled conditions, i.e., the context in our design was largely controlled. Most context elements were fixed during the experiment, except the knowledge and experience of the architect. To achieve comparable results, we asked the participants to provide us with their experience in assessing business cases, the number of years work experience, and the number of years as an architect.

The checklist and the business case were tested in three different pre-tests before we started the experiment. The goal of the first pre-test was to validate the usefulness and content of the checklist with two experienced enterprise architects. Based on this test we changed the order of checkpoints, the language from English to Dutch, and made some minor content adjustments. We also got confirmation that a checklist can support enterprise architects in systematically improving a business case. In the second pre-test we asked two experienced enterprise architects to test the usefulness of the business case. The outcome was positive; the business case in question was judged as leaving enough room for improvement. Another advantage of this business case is that most architects have at least a basic understanding of the context, i.e., Dutch government. The second test led to some minor changes in the checklist. In the third pre-test, again with two architects, we tested how much time was needed for the experiment. We conducted the trial exactly the same way as the actual experiment. We found a maximum of 90 minutes acceptable. It took the two architects 66 and 81 minutes respectively. Both architects thought that the checklist would have added value for architects in assessing business cases.

D. Focus groups

After each experiment, participants were asked to provide feedback on the experiment. The different groups thus also acted as confirmatory focus groups for artifact refinement and evaluation [33]. We asked the participants with a checklist whether they used it and to what extent. We also asked for their ideas how to successfully use the checklist. Furthermore, the participants of the 3rd, 4th, and 5th group were provided with preliminary results of the previous groups and asked for their thoughts on those results. One of the researchers took notes during the focus group session.

IV. Execution

A. Data Collection and Analysis

All architects in our experiment were provided with the same business case. The assignment of subjects to each treatment was selected randomly. Half of one group of architects was asked to assess the business case using a checklist and the other half without a checklist. Participants were not aware there was going to be a checklist, nor that some of their fellow participants were using one. Architects were given the same assignment; developing improvement suggestions for the business case. The precise question that was asked: “please formulate improvement suggestions for this business case so that a good decision can be made”. The participants with a checklist were asked to read the checklist in advance and use it to assess the business case. They were instructed to use the checklist however they saw fit.

For the introduction of the experiment to the participants and for the collection of data, different forms were created. The assignment was introduced by means of an instruction letter. Copies of the business case and forms were made available to all subjects and were put in one envelope. The envelopes with a checklist had a different introduction letter compared to the envelopes without the checklist. The envelopes were numbered and distributed randomly across the participants. Every participant was given one envelop. The differences between envelopes were not known to the participants. The participants had to bring their laptops so that data could be collected at the experiment location. By the end of the experiment the participants had to email their improvement suggestions to one of the researchers. Participants had also to sign a consent form and to fill in a demographics form. By means of the demographics form we collected data on their experience in assessing business cases, their working experience, and their experience as an architect.

The five experiments were held in November and December 2018, lasted about 3 hours per experiment and were supervised by two researchers. The experiment started with a short introduction of the participants and the experiment. Subsequently, the envelopes were randomly distributed among the participants and the experiment began. Participants were asked to work individually and not to share their thoughts with others. It took the participants between 60 and 105 minutes to complete the task. After a break the experiment was discussed with the participants and only then they became aware of the different treatments. During the experiments we collected improvement suggestions for the business case. We received in total 365 improvement suggestions. Eight suggestions were excluded because the same participant came with identical suggestions or the suggestion did not indicate any improvement or shortcoming. One participant was excluded because of a lack of any architecture experience. Accordingly, another five suggestions were excluded. In the end we included 352 improvement suggestions in the analysis. These suggestions were linked to keywords. Keywording is meant to assist in comparing the data and to remove too much individuality from the extracted data [34]. Using keywords, we were able to check whether the improvement suggestions could be related to one of the checkpoints on the checklist. The list of keywords was initially created based on the pre-tests and contained 16 keywords. After each experiment the list of keywords grew. In the end we had 22 different keywords. Each improvement suggestion has at least one keyword. The keywords were assigned by two researchers independently from each other. Differences were discussed and resolved. The total number of keywords assigned to the improvement suggestions is 415. By the end of the experiments all 352 improved suggestions were assessed by five experts. Each of

them gave each suggestion a score from 1 to 5. The experts worked independently from each other. They did not know whether an improvement suggestion was the result of one of the two treatments. The experts were asked to use their own frame of reference when assessing the improvement suggestions. Statistical analysis was conducted with SPSS, a software package for statistical analysis. Researchers took field notes throughout the experiments.

TABLE I. CHECKLIST

Checkpoint	Underlying questions
Check the fit with the business strategy	To what extent does this business case realize the business strategy? Is this business case in line with the business strategy? Is this business case in conflict with the business strategy? To what extent does this business case meet the set business objectives?
Check the future options	To what extent does this business case generate opportunities that can be redeemed in the future?
Check the stakeholders' concerns	To what extent are the concerns of the stakeholders known? Are these concerns sufficiently reflected in the business case?
Check the solution alternatives	What are the solution alternatives to realize this business case? Are these solution alternatives recognized in the business case? Are the solution alternatives weighed up? Is the proposed solution motivated? Does the proposed solution meet the interests of the stakeholders? Does the proposed solution make it possible to redeem the benefits of this business case?
Check the relationships with other business cases	Can other business cases benefit from this business case? Can this business case ruin other business cases?
Check the consequences for the current state	What are the consequences for the current landscape? Are these consequences recognized in the business case? To what extent can this business case ruin the current landscape?
Check the fit with the future state architecture	To what extent does this business case realize the future state architecture? Is this business case in line with the future state architecture? Is this business case in conflict with the future state architecture? Is this business case in line with architecture principles, policies and standards? Is this business case in line with current market developments?
Check the feasibility	To what extent is this business case feasible? What makes it difficult to realize the solution for this business case? What makes it difficult to realize the benefits of this business case? Is the feasibility of the solution addressed in the business case? Is the feasibility of the benefits addressed in the business case?
Check the risks	What are the main risks to realize this business case? Are these risks recognized in the business case?

B. Checklist

Table I contains the checklist that we used in this research. This checklist underwent seven iterations. We first discussed the checklist among researchers and modified it. Then three pre-tests were conducted in which we gathered feedback and modified the checklist. One of the discussions we had during the development of the checklist was whether to include benefits and costs as checkpoints. The participants of the first pre-test were strongly against the introduction of benefits and costs. They found that architects should mainly concern themselves with the content and solution alternatives, and not with the benefits and costs. So we decided not to include benefits and costs as checkpoints.

V. RESULTS

A. Descriptives

Table II contains demographic data and demonstrates that the participants with checklist are on average older and have more experience. Five architects had no experience in assessing business cases and 26 did have experience. Five of the 32 participants have the Belgian and 27 the Dutch nationality. Four of the 32 participants were women, the rest men. All participants have experience with architecture, but have different job names like enterprise architect, architect, IT architect, business architect, domain architect, solution architect, tribe architect, consultant, program coordinator, business engineer, CEO, and team lead.

The first group consisted of six participants from an IT service provider. The second group had three participants from a consultancy company. The third group comprised 11 participants from the CIO-platform, a Dutch network organization for CIOs. The fourth group had five participants from a Belgian consultancy company and the fifth group consisted of seven participants from the NAF, a Dutch architecture network.

Five experts rated all improvement suggestions. These experts have extensive experience in the development and assessment of business cases. The average number of years working experience is 24. The average age of the experts is 48.4.

TABLE II. DESCRIPTIVES PER TREATMENT AND OVERALL

	Treatment one	Treatment two	All participants
Number of participants	17	15	32
Experience in assessing business cases	14	12	26
No experience in assessing business cases	2	3	5
Unknown experience in assessing business cases	1	0	1
Average age	50.12	43.06	48.16
Average number of years working experience	24.88	20.87	23.00
Average number of years experience with architecture	13.41	11.30	12.42

TABLE III. RATINGS AND CRITERIA PER EXPERT

Expert	Mean of Treatment One	Mean of Treatment Two	Role	Criteria
1	2.74	3.03	Senior manager	• My main criterion was: would I take up this suggestion and improve something concrete with it, if this had been my business case? • The main reasoning was whether the suggestion really adds something to understanding, decision making and/or benefit tracking. (e.g., clarification, simplification, tightening). • The key question, in my opinion, is "Why should we do this"?
2	3.14	3.02	Senior adviser	• The degree of impact the suggestion has on the following aspects: benefits, context, content, substantiation of the advice, editing of the document, scenarios, future state, implementation aspects.
3	2.37	2.41	Program manager	• Is it an improvement because thought has been given to future maintainability, manageability, usability that requires IT experience? These comments are very valuable and here is the added value of engaging the enterprise architect. These comments can vary in the ranking between 1 and 5. • Is it an improvement that concerns the description of business benefits, business strategy, calculation, etc.? These are essential remarks that require that someone is experienced on the business process in question. This is not unique for the enterprise architect (a good controller could have done it) but it does add value. These remarks are not higher than 3 in the ranking. • Is it an improvement linguistically, semantically, or about the document structure etcetera? Every reader can do this. The enterprise architect has no unique added value. The comments can be ok, but they do not exceed 3 in my ranking.
4	3.40	3.66	CIO	• The degree of argumentation and constructiveness of the suggestion. • The extent with which I could agree with the content of the suggestion.
5	3.10	3.35	Head of CIO Office	• I always look broadly and start from projects that are initiated from a business perspective. In other words, for me the whole process is important. This process covers the business idea, IT realization, the implementation of the process and all IT and non-IT tools, and includes maintenance, management and life cycle. I started working from this perspective.

B. Experiment and Focus groups

The ratings of the experts, their job titles as well as the criteria that they used are shown in Table III. Table III demonstrates that four out of five experts gave the improvement suggestions from participants without a checklist a higher score than the improvement suggestions of participants with a checklist. Table III also shows that although the experts used different criteria in the assessment of the improvement suggestions, they are fairly close in their judgement. Table IV contains the main results of the experiment and shows that the mean of the scores of the participants without checklist ($\mu\ S_{withoutchecklist} = 3.07$) is higher than the mean of the scores of the participants with checklist ($\mu\ S_{checklist} = 2.97$).

We conducted an independent sample test (t-test) in SPSS with the scores of the participants as the dependent variable and the treatment as the independent variable to determine whether the differences between the two treatments are significant. The scores per treatment are normally distributed and homogeneous. The two assumptions of the t-test (normality and homogeneity) are thus met. The result of the t-test demonstrates that the differences between the mean values of the two treatments are not significant, $t = .81$, $df = 30$ and $p = 0.42$. This means that for assessing the business case, there is no significant difference between a participant using the checklist or not [35]. H0 is thus accepted and H1 rejected.

Seven of the 14 keywords in Table V can be directly related to checkpoints on the checklist (with asterisk). Seven other keywords cannot.

TABLE IV. STATISTICS PER TREATMENT AND OVERALL

	Treatment one	Treatment two	Total
Number of participants (N)	17	15	32
Number of improvement suggestions	195	157	352
Average number of improvement suggestions per participant	11.47	10.47	11.00
Mean of scores of improvement suggestions per participant	2.97	3.07	3.02
Standard deviation of scores of improvement suggestions per participant	0.34	0.41	0.37

The analysis of keywords shows that the highest scoring improvement suggestions relate to a checkpoint that cannot be extracted directly from the checklist, namely "Benefits" and "Objectives". What stands out in Table V, is that the improvement suggestions of participants with a checklist (treatment one) refer much more often to checklist related keywords than improvement suggestions of participants without a checklist (treatment two). Participants without a checklist came up with more improvement suggestions about the readability of the business case than participants with a checklist.

Of the nine checkpoints on the checklist, two are not included in Table V because they have been linked as a keyword to improvement suggestions less than 10 times. These are listed in table VI. Table VI demonstrates that "Feasibility" has the highest score of all keywords. On the other hand, it is only linked seven times to improvement suggestions.

TABLE V. KEYWORDS RANKED ACCORDING TO THE SCORES OF IMPROVEMENT SUGGESTIONS TO WHICH THEY ARE LINKED

Keyword	Mean of score	Total number of key-words	Linked to improvement suggestions treatment one	Linked to improvement suggestions treatment two
Benefits	3.41	49	21	28
Objectives	3.33	31	11	20
Business vision & strategy*)	3.31	11	10	1
Risks*)	3.28	16	11	5
Alternatives*)	3.25	39	31	8
Current state*)	3.17	12	10	2
Requirements	3.16	16	7	9
Stakeholders*)	3.06	33	22	11
Context*)	3.05	15	13	2
Costs	2.99	38	22	16
Approach	2.91	10	3	7
Future state*)	2.82	20	14	6
Functionalities	2.82	12	9	3
Readability	2.42	66	25	41

TABLE VI. REMAINING KEYWORDS BELONGING TO CHECKPOINTS ON THE CHECKLIST

Keyword	Mean of score	Total number of keywords	Linked to improvement suggestions treatment one	Linked to improvement suggestions treatment two
Feasibility*)	3.46	7	6	1
Future options*)	3.16	5	5	0

We also learn from Tables V and VI that the keywords that refer to the checklist, are much more common among participants with a checklist than participants without a checklist. The checklist does its job. Apparently, participants with a checklist are highly influenced by the checkpoints on the checklist and refer to these checkpoints much more often compared to participants without a checklist. The checklist is thus valuable because it decreases the risk of overlooking insights.

A possible confounding factor is the experience of the architect in assessing business cases. The mean of scores of improvement suggestions of all 26 experienced architects is 3.00. The mean of scores of all 5 inexperienced architects is 3.16. The architects with no experience in assessing business cases developed slightly better improvement suggestions as compared to architects with experience in assessing business cases. Experience seems of no relevance. A correlation analysis in SPPS between the mean score of the participant and the number of years of experience as architect confirms this result. The Pearson's correlation coefficient $r = .01$ and $p = .97$, showing that the relationship between experience as architect and score is very weak. The correlation between the mean score of the participant and the number of years of working experience is $r = .01$ and $p = .94$, demonstrating that also the relationship between experience and score is very weak. It should be noted that we have a relatively small sample size (N=32).

VI. DISCUSSION AND CONTRIBUTION

A. Discussion

The results of this research are surprising. We expected participants with a checklist to make better improvement suggestions compared to participants without a checklist. All the more so because in the discussion with participants after the experiment, they almost unanimously indicated that the checklist has added value. In addition, the participants in treatment one are on average older, have more work experience, and are more experienced as architects. Why is there no difference, and why is the average score of the participants with a checklist even lower?

The discussions at the end of the experiment revealed that some participants who had a checklist at their disposal did not use it at all or only glanced at it. One of the participants explained: "I read the checklist and thought, yes, I know all of this, and after that I did not explicitly use the checklist anymore". Another participant with a checklist first assessed the business case without a checklist and then supplemented the improvement suggestions based on the checklist. According to his own words, he covered 90% without using the checklist. However, looking at his submission, it appears that he paid attention to only 6 of the 9 checkpoints (66%). Gawande argues that "it feels somehow beneath us to use a checklist, an embarrassment" [25]. Apparently, some participants felt embarrassed too. They tend to prefer their own way of working and frame of reference over a standard operating procedure and checklist. A checklist may feel like additional bureaucracy and block the initiative, intuition, and common sense. The SLR suggests that positive results of the introduction of a checklist are associated with behaviour, implementation, and communication. When we want architects to make good use of a checklist we need to properly introduce and communicate this in advance.

Looking at the submissions of treatment one, some included their own frame of reference, in terms of setting out a framework, a set of key performance indicators, or a list of what they thought was important. Just as if they created their own checklist at the moment they discovered that they were supposed to use one in the experiment. Gawande points to issues professionals have in using a checklist. One of these issues is that individual autonomy seems the ideal we should aim for. Novelty and excitement is what we like. Gawande argues that in situations of high stakes and complexity protocols and procedures are required, and in line with that, discipline. But discipline is hard and something we have to work at [25]. Using a pre-defined checklist like ours, also requires discipline.

A closer investigation of the improvement suggestions reveals that some participants with a checklist followed the checklist closely but came up with rather meaningless suggestions. In the NAF group it was rightly noted that instructions are also needed for using the checklist. This is to prevent people from using the checklist indiscriminately, as has indeed happened in some cases. On average, participants without a checklist came up with fewer improvement suggestions compared to participants with a checklist (10.47

compared to 11.47), but these suggestions are of a higher quality. Apparently, without a checklist, one is forced to find out where to assess the business case. This may take more time, resulting in fewer suggestions, but better thought through. This also indicates that the way the checklist is used determines the success of the assessment of a business case. A form of instruction in how to use the checklist and how to interpret the concepts behind the checkpoints is required.

Another reason why the checklist did not lead to higher quality improvement suggestions may lie in its completeness. If "Benefits" had been included as a separate checkpoint on the checklist, participants with a checklist would probably have named it more often and would have scored higher. The same goes for "Objectives". As can be seen in the checklist, "Objectives" is not a separate checkpoint, but part of "Business vision and strategy". The NAF group made a plea for a separate checkpoint. One of the discussions with the groups was whether the checklist should also include aspects that should be assessed by other disciplines involved in a business case. The most sensitive point is whether the architect should also assess costs and benefits. Some architects, as we found in our pre-test, are opposed to this. After all, you have controllers for costs and benefits. Others, like the CIO-platform, the Belgian consultancy company, and the NAF groups, are strongly in favour of the architect taking the costs and benefits into account, because solution alternatives have budgetary constraints, have an impact on costs and benefits, and many decision-makers talk in terms of costs and benefits. If you want to play along as an architect, you cannot avoid translating your solutions into costs and benefits. One of the participants with a checklist argued: "I did not use the checklist, because a business case is about benefits and costs, and these are missing from the checklist". The majority of the participants in the focus groups are of the opinion that costs and benefits should be added to the checklist as checkpoints. We are aware that the checklist then may contain checkpoints that are not unique for an enterprise architect. Like expert 2 noted, some checkpoints (future options, solution alternatives, relationships with other business cases, consequences for the current state, fit with the future state architecture, and feasibility) are specifically attributable to enterprise architects. Other checkpoints (fit with business strategy, stakeholder concerns, risks, costs and benefits) require the attention of different stakeholder groups including enterprise architects. By introducing new checkpoints, the checklist could become less EA specific, since the new checkpoints are not specifically attributable to enterprise architects. We suggest making the checkpoints as EA-specific as possible. Instead of asking "To what extent are the benefits of this business case reliably estimated?", we should question "To what extent do the solution alternatives contribute to the benefits?". In the end, the completeness of the checklist may be another success factor in using a checklist to assess a business case.

Domain knowledge is another success factor. Take for example "Feasibility". As a keyword, it is only mentioned seven times. A reason may be that the business case that we used in the experiment is not easy to understand without domain knowledge. For example, knowledge on what made previous projects successful in the same context? Participants without a checklist did not even mention checkpoints that were difficult to answer. They just don't think about it. Domain knowledge is also required for checkpoints like "Check the fit with the future state architecture" and "Check for future options". When we look closer at the improvement suggestions with the keywords "Future State Architecture" and "Future Options", these suggestions were regularly along the lines of "not seen" or "do not occur"; quite meaningless. The lack of domain knowledge among the participants with a checklist, makes it difficult for them to translate checkpoints into good improvement suggestions. In a real-life situation, the use of the checklist in combination with domain knowledge almost certainly leads to higher quality improvement suggestions compared to not using a checklist.

One of the results of this research is that participants with a checklist refer much more often to checklist related keywords than participants without a checklist. One of the participants, who used the checklist completely, explained: "this checklist ensures that your mindset will be focused in advance when you start reading the business case. The checklist provides structure. You start reading and assessing the business case with a kind of frame in your head". This can be one of the reasons that participants with checklists had fewer improvement suggestions on the readability of the business case compared to participants without a checklist. It was also remarked that "the checklist contains checkpoints you do not immediately think of, like the relationship with other business cases". In any case, by using the checklist completely, the chance of overlooking valuable insights decreases.

Our checklist is not yet finished. It should be modified according to the results of this research and tested in a real-life situation, i.e., a situation where the architect has domain knowledge. The best test is where the architect uses the checklist as a do-confirm checklist. This means that the architect first assesses the business case from his/her own perspective and then uses the checklist to see if there are any checkpoints that were overlooked and should be taken into account. By doing this in a number of real-life situations, the checklist can be further refined. Another test is to compare the assessment results of the architect with the assessment results of other stakeholders. This comparison may also lead to further refinement.

The image that emerges from the experiment and the discussion is that of architects who, over time, developed their own frame of reference and work accordingly. This is in line with Hoorn et al. who concluded that "architects are rather complicated individuals", and "with respect to sharing architectural knowledge, architects are not the 'community builders' many expect them to be. Instead, they are rather lonesome decision-makers who act in splendid isolation" [36]. In the end, the objective should be to create one checklist for all disciplines involved in a business case. According to Gawande, teamwork is essential [25]. The development of a business case is teamwork. It would be great to have one checklist for a business case covering all

aspects of all disciplines, including EA, and discuss these aspects among all disciplines rather than assess the business case per discipline. Such a checklist can really function as a boundary object between different disciplines.

What came as a surprise in this study, is that the experience of architects did not matter in the assessment of a business case. Strano and Rehmani argue that one of the competencies that is most needed to be effective in the role of enterprise architect is business acumen [37]. Business acumen refers, amongst other things, to managerial qualities such as experience with preparing business cases. Our research demonstrates that this form of business acumen is not necessarily based on experience. Op't Land et al. argue that experience is one of the ways to acquire the competencies of an enterprise architect, while education is another [3]. They also notice that job ads for enterprise architects typically ask for more than 5 years of experience [3]. Our study demonstrates that this is not necessary for all EA tasks. The notion that experience is of less relevance for an enterprise architect requires further research.

The use of a checklist as such does not help architects to improve the quality of a business base. However, the chance of success can be increased when:

- the checklist is properly introduced and communicated;
- the checklist is part of a procedure that is strictly followed by architects;
- the user of the checklist is educated in how to use it and the concepts behind the checkpoints;
- the user has knowledge of the domain of the business case;
- the checklist is more complete and is extended with checkpoints on benefits, costs, and objectives.

Under these conditions the checklist most likely improves the quality of improvement suggestions in the assessment of a business case as compared to not using a checklist. In any case, the risk of overlooking valuable insights is reduced. However, one point of concern is the image of an architect as a lonesome decision-maker. In the improvement of the quality of a business case, the architect should at least be a team player, and preferably a team builder. One who understands how to interact with other disciplines involved in the development of a business case and is even able to lead the process to develop a world class business case.

B. *Contribution*

The main contribution of this study is that we developed a theory for design and action [9]. We built and evaluated a checklist that can be used by enterprise architects in their assessment of a business case.

For EA practitioners, this checklist can be used in the development and assessment of a business case. However, what is even more important for EA practitioners, is the insight that a checklist as such does not make the difference. It is a combination of different factors that determines the success of using a checklist.

For researchers, this study adds to the EA knowledge base with a checklist and insights on how to use it. Next to that, opportunities for future research were revealed. The checklist should be revised and tested in case studies. Another interesting area for future research is how much experience is needed to function as an enterprise architect.

VII. THREATS TO VALIDITY AND LIMITATIONS

The following threats may occur to the validity of an experiment: conclusion, internal, construct, and external validity [31]. Conclusion validity was achieved by statistical tests, preparation, and measurement of results. A weak point in our design is that the quality of the improvement suggestions is measured by experts, which is a subjective way of measuring. To reduce the subjectivity we included five experts in the scoring of the improvement suggestions. Another weak point is the sample size. We had 32 participants, which is too small to draw statistically relevant conclusions. Due to the small sample size, the division of participants over treatments was a little imbalanced. Participants of treatment one are on average older and have more experience. Internal validity was achieved by randomly dividing subjects over treatments. Construct validity was achieved by using a real-life business case and by pre-testing this business case to make sure it contains room for improvement. Another measure was to ask experts to rank the improvement suggestions. This is also a weak point in our design, because it is a subjective way of measuring. We are aware that some possible biases may occur, such as hypothesis guessing and evaluation apprehension [30]. External validity was achieved by conducting the experiment at five different locations, each with a different group of architects, and by using a real-life object.

A limitation of this research is that different kinds of architects participated in the experiment. Results may be biased according to the background and the belief system of the architect. Another limitation is that we used one similar business case to achieve comparable results among participants. Another business case may generate different results. Further research is required to evaluate and refine the checklist on different business cases. A third limitation is that we focused on the evaluation of the checklist. A more profound design of the checklist may yield different results. Consortium research is an approach to further strengthen the design of the checklist [38].

VIII. CONCLUSION

The aim of this study was to develop and evaluate a checklist by which enterprise architects can assess a business case and thus contribute to the quality of a business case. This should ultimately improve the quality of the IT investment decision. We started with a systematic literature review on checklists, followed by the development of a checklist and the evaluation of the checklist by means of an experiment and discussion in focus groups. The results of our research demonstrate that a checklist alone does not make any difference in the quality of the improvement suggestions of architects that assessed a business case. Our research also reveals preconditions for the use of the

checklist. First, the checklist should be properly introduced and communicated. Second, the checklist should be part of a procedure that should be followed strictly by architects. Third, the user of the checklist should be trained in how to use the checklist and the concepts behind the checkpoints. Fourth, the user should have knowledge of the domain of the business case. Fifth, the checklist should be extended with checkpoints on benefits, costs, and objectives. When these factors are applied, the use of the checklist is likely to improve the quality of a business case, though further research is needed to confirm this. Another surprising finding from this research is that experience makes no difference in the assessment of a business case by architects. This is also an area for further research. In any case, the use of the checklist reduces the risk of overlooking valuable insights that would otherwise have gone unnoticed.

REFERENCES

[1] P. Weill and J. Ross, IT Governance, How Top Performers Manage IT Decision Rights for Superior Results, Boston: Harvard Business School Press, 2004.

[2] Standish Group, The CHAOS Summary 2015 Report, Boston: The Standish Group International, 2015.

[3] M. Op't Land, E. Proper, M. Waage, J. Cloo, and C. Steghuis, Enterprise architecture: creating value by informed governance, Berlin: Springer Science & Business Media, 2008.

[4] E. Niemi and S. Pekkola, "Enterprise Architecture Benefit Realization: Review of the Models and a Case Study of a Public Organization," ACM SIGMIS Database, vol. 47, no. 3, 2016, pp. 55-80.

[5] R. Foorthuis, M. van Steenbergen, N. Mushkudiani, W. Bruls, S. Brinkkemper, and R. Bos, "On Course, but not There Yet: Enterprise Architecture Conformance and Benefits in Systems Development," International Conference on Information Systems, paper 110, 2010.

[6] M. van Steenbergen and S. Brinkkemper, "Modeling the contribution of enterprise architecture practice to the achievement of business goals," Information Systems Development, Boston: Springer, 2009, pp. 609-618.

[7] M. Lange, J. Mendling, and J. Recker, "An empirical analysis of the factors and measures of Enterprise Architecture Management success," European Journal of Information Systems, vol. 25, no. 5, 2016, pp. 411-431.

[8] M. van den Berg, R. Slot, M. van Steenbergen, P. Faasse, and H. van Vliet, "How Enterprise Architecture Improves the Quality of IT Investment Decisions," Journal of Systems and Software, vol. 152, 2019, pp. 134-150.

[9] S. Gregor, "The nature of theory in information systems," MIS quarterly, 2006, pp. 611-642.

[10] P. Johnson, M. Ekstedt, E. Silva, and L. Plazaola, "Using Enterprise Architecture for CIO Decision-making: on the Importance of Theory," Proc. Second Annual Conference on Systems Engineering Research, 2004, paper 116.

[11] M. Janssen, "Sociopolitical aspects of interoperability and enterprise architecture in e-government," Social Science Computer Review, vol. 30, no. 1, 2012, pp. 24-36.

[12] T. Tamm, P. Seddon, G. Shanks, and P. Reynolds, "How does enterprise architecture add value to organisations," Communications of the Association for Information Systems, vol. 28, no. 1, 2011, pp. 141-168.

[13] V. Boucharas, M. van Steenbergen, S. Jansen, and S. Brinkkemper, "The contribution of enterprise architecture to the achievement of organizational goals: a review of the evidence," Trends in Enterprise Architecture Research, Berlin: Springer, 2010, pp. 1-15.

[14] The Open Group, TOGAF™ Version 9, Van Haren, 2009.

[15] M. Blosch and B. Burton, Future of Enterprise Architecture 2025: Evolving From Enterprise to Ecosystem, Gartner Research Note G00269850, 2014.

[16] J. Gøtze, "The changing role of the enterprise architect," Proc. IEEE Intern. Enterprise Distributed Object Computing Conference Workshops, 2013, pp. 319-326.

[17] CIO Council, Clinger-Cohen Core Competencies & Learning Objectives, 2012.

[18] R. Wagter, M. van den Berg, J. Luijpers, and M. van Steenbergen, Dynamic enterprise architecture: how to make it work. Hoboken: John Wiley & Sons, 2005.

[19] J. Lapalme, "Three schools of thought on enterprise architecture," IT professional, vol. 14, no. 6, 2012, pp. 37-43.

[20] D. Greefhorst and E. Proper, Architecture Principles, The Cornerstones of Enterprise Architecture, Berlin: Springer, 2011.

[21] ISACA, The Business Case Guide: Using Val IT 2.0, 2010.

[22] J. Ward, E. Daniel, and J. Peppard, "Building a better business case for IT Investments," MIS Quarterly Executive, vol. 7, no. 1, 2008, pp. 1-14.

[23] E. Berghout and C. Tan, "Understanding the impact of business cases on IT investment decisions: An analysis of municipal e-government projects," Information & management, vol. 50, no. 7, 2013, pp. 489-506.

[24] C. Bentley, Prince2: a practical handbook, Routledge, 2012.

[25] A. Gawande, The checklist manifesto, Penguin Books India, 2011.

[26] D. Kahneman, Thinking, fast and slow, London: Allen Lane, 2011.

[27] J. Dean Jr. and M. Sharfman, "Does decision process matter? A study of strategic decision-making effectiveness," Academy of management journal, vol. 39, no. 2, 1996, pp. 368-392.

[28] A. Hevner, S. March, J. Park, and S. Ram, "Design science in information systems research," Management Information Systems Quarterly, vol. 28, no. 1, 2004, pp. 6.

[29] S. Star and J. Griesemer, "Institutional ecology, translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39," Social studies of science, vol. 19, no. 3, 1989, pp. 387-420.

[30] D. Kahneman, D. Lovallo, and O. Sibony, "Before you make that big decision," Harvard business review, vol. 89, no. 6, 2011, pp. 50-60.

[31] C. Wohlin, P. Runeson, M. Höst, M. Ohlsson, B. Regnell, and A. Wesslén, Experimentation in software engineering. Heidelberg: Springer Science & Business Media, 2012.

[32] Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. Maatschappelijke Business Case Stelselcatalogus releases 3 en 4 v1.0, 2016. Last accessed on January 28, 2019, Retrieved from https://www.gemmaonline.nl/index.php/Bestand:Actie_20160525_127_bijl_16_Business_Case_Stelselcatalogus_releases_3_en_4_v1_0.pdf.

[33] M. Tremblay, A. Hevner, and D. Berndt, "The use of focus groups in design science research," Design Research in Information Systems, 2010, pp. 121-143.

[34] K. Petersen, S. Vakkalanka, and L. Kuzniarz, "Guidelines for conducting systematic mapping studies in software engineering: An update," Information and Software Technology, vol. 64, 2015, pp. 1-18.

[35] A. Field, Discovering statistics using SPSS. Sage publications, 2009.

[36] J. Hoorn, R. Farenhorst, P. Lago, and H. van Vliet, "The lonesome architect," Journal of Systems and Software, vol. 84, no. 9, 2011, pp. 1424-1435.

[37] C. Strano and Q. Rehmani, "The role of the enterprise architect," Information systems and e-business management, vol. 5, no. 4, 2007, pp. 379-396.

[38] H. Österle and B. Otto, "Consortium research", Business & Information Systems Engineering, vol. 2, no. 5, 2010, pp. 283-293.