Association for Information Systems
AIS Electronic Library (AISeL)

ECIS 2012 Proceedings

European Conference on Information Systems (ECIS)

5-2-2012

SUPPORTING DYNAMIC REUSE IN BUSINESS CASE DEVELOPMENT

Bart-Jan van Putten

Thomas Irrenhauser

Theo Dirk Meijler

SUPPORTING DYNAMIC REUSE IN BUSINESS CASE DEVELOPMENT[1]

Van Putten, Bart-Jan, Humboldt-Universität zu Berlin, Institute of Information Systems, Spandauer Straße 1, 10178 Berlin, Germany; SAP Research Dresden, Chemnitzer Straße 48, 01187 Dresden, Germany, bart-jan.van.putten@sap.com

Irrenhauser, Thomas, Technische Universität München, Institut für Werkzeugmaschinen und Betriebswissenschaften, Boltzmannstraße 15, 85748 Garching, Germany, thomas.irrenhauser@iwb.tum.de

Meijler, Theo Dirk, SAP Research Dresden, Chemnitzer Straße 48, 01187 Dresden, Germany, theo.dirk.meijler@sap.com

Abstract

Business case development (BCD) is a complex activity, which can potentially be improved by supporting the reuse of investment criteria and valuation methods. The goal of this research was to improve the usefulness and usability of business case frameworks (BCFs), while limiting the effort required to develop and maintain static databases of reusable components. Therefore, an approach was proposed for the dynamic reuse of business case components and contrasted with static reuse of business case components. In the dynamic approach, the reusable, domain-specific criteria and methods do not need to be pre-defined by experts in templates and taxonomies, but can be reused from earlier business cases. To test whether support for dynamic reuse improves BCFs, a usability experiment was set up. Three types of support for the reuse of criteria were compared: (1) recommendations, based on collaborative filtering and representative for the dynamic approach, (2) templates, representative for the static approach, and (3) no support. The task represented a simplified BCD activity and was completed by 208 people. The main results show that although the recommendations are as effective as the templates, they are the preferred type of support.

Keywords: Business case, Cost-benefit analysis, Value, IS, Investment, Evaluation, Reuse.

[1] This research is partly funded by the German Ministry of Economics and Technology (BMWi) under contract 01NA10018, project RAN, from February 2011 through February 2012. The authors would like to thank the following people for their support during the preparation of the experiment: Andrea Lehmann, Angelika Salmen, Antonio Krüger, Bettina Laugwitz, Elke Haenle, Hermann-Josef Schwab, Hervé Couturier, Jelke Bethlehem, Jochen Rode, Nadja Schubert, Oliver Günther, Sai Prasanna Naidu Byyarapu, Sebastian Reuther, Sven Horn. Moreover, we are grateful to all people who participated in the experiment.

1 Introduction

For decades, determining the value and success of information systems (IS) has been high on the agendas of researchers, cf. (King & Schrems 1978; F. D. Davis 1989; Urbach et al. 2009; Schryen 2010). Many IS investments are strategic in nature, have long-term, hard to quantify benefits and incur indirect costs (Irani 2002). 68% of IS projects cope with problems when having to identify the potential benefits of the investment, while 85% have problems with their quantification (Ballantine & Stray 1999). This may lead to investment in the wrong projects, over-investment, under-investment, and makes it attractive to creatively adjust estimates, see e.g. (Ward et al. 2008).

To support IS investment evaluation (ISE), in most organisations some kind of business case (BC) is developed. A BC is *"a recommendation to decision makers to take a particular course of action for the organisation, supported by an analysis of benefits, costs and risks"* (Gambles 2009). Business case development (BCD) is the process of realising the BC as an 'artefact', by gathering and analysing data to define and valuate evaluation *criteria* (Bacon 1992), presenting it in documents, spreadsheets or presentations. It may take place before project execution for investment appraisal, during project execution for monitoring and control, or after project execution for organizational learning (Van Putten et al. 2011).

When trying to estimate the values of the criteria selected for use in the new BC, a BC developer may spend days on defining *methods*, i.e., ways to put a qualitative, quantitative non-financial, or financial value on these criteria, see e.g., (Renkema & Berghout 1997). This is especially hard for the intangible benefits and indirect costs (Irani 2002). Instead, criteria and methods may also come from databases which are specifically designed for the purpose of reuse, e.g., templates and taxonomies of criteria and methods with browse and query functionality for their retrieval (Irani et al. 2006).

BCD can be supported by a business case framework (BCF), which often comes in the form of a spreadsheet template with some pre-defined criteria and methods. Today's BCFs are, however, often too generic, providing little support for BC developers who need to define domain-specific criteria and methods, e.g., in the domain of supply chain management, which will be used as an example throughout this paper. Other times, BCFs are sufficiently domain-specific, but are based on templates and taxonomies, which are explicitly defined by domain experts. Such BCFs are expensive to develop and maintain and are therefore often limited to one domain.

This paper proposes a new mechanism applicable in BCFs, namely *dynamic reuse of BC components*. This in contrast to the *static reuse of BC components* as this is known from templates and taxonomies. Dynamic reuse implies that components such as criteria and methods that were developed and used for earlier BCs, when stored in a structured manner, can be reused in later BCs on the basis of a recommender algorithm. This paper focuses on how dynamic criteria reuse could work and describes the result of a large-scale usability experiment, comparing different types of support for reuse, namely no support, templates (static reuse) and recommendations (dynamic reuse).

The following research questions are addressed in this paper:

RQ1. How can support for reuse improve the usefulness and usability of business case frameworks, while limiting the effort required to develop and maintain static databases of reusable components?

RQ2. To what extent does support for the dynamic reuse of criteria (being the focus of this paper), improve the usefulness and usability of business case frameworks?

This paper is further structured as follows: In section 2 we describe a scenario from the domain of supply chain management, motivating why it is helpful to be able to dynamically reuse the criteria from an earlier BC in a new BC. In section 3 we investigate existing BCFs, further substantiating the shortcomings mentioned above. In section 4 we detail the dynamic reuse concept. Section 5 explains the research method, i.e., the usability experiment, and section 6 presents the results. Finally, section 7 concludes this paper.

2 Motivation for the Reuse of BC Components

The following scenario illustrates how reuse could take place during BCD:

Imagine an earlier BC on the use of barcodes for tracking items in the supply chain. Due to the use of barcodes, it may become easier for warehouse workers to have fairly precise data on stock levels, which may in turn make it possible to decrease those stock levels and thus decrease warehousing costs. Criteria in this BC may be 'ease of identification', 'stock level' and 'warehousing costs'. Some years later, when developing a BC for the use of radio-frequency identification (RFID), the BC developer would again need to decide which criteria to use. Rather than defining all criteria from scratch, the developer may reuse criteria from earlier BCs. In this case, the criteria 'ease of identification', 'stock level' and 'warehousing costs' would probably also be relevant for the BC for RFID.

Support for reuse has the potential to make life easier for BC developers, by easing the identification of criteria and the development of methods. Estimations in BCs may become more reliable, because the BC developer can build upon the methods developed by others. Moreover, estimations may be compared against benchmarks, e.g., aggregated values from earlier BCs. Reuse also enables longer-term BC development and monitoring and allows BC developers to get skilled at using certain methods. Finally, when BCs are structured similarly, BC *evaluators* may find it easier to compare BCs and decide how to invest.

3 Limitations of Business Case Frameworks

For this research, from the perspective of reuse, two classes of BCFs are distinguished (Table 1):

- Traditional BCFs: Mainly support the development of the BC in terms of its cash flow, often accompanied by metrics such as Return-on-Investment (ROI), Net Present Value (NPV), or Internal Rate of Return (IRR). These BCFs are often domain agnostic and have rigid structures. Sometimes, new criteria can be entered, but those then need to be quantified directly in financial terms. These BCFs are often implemented as a stand-alone Microsoft Excel file and do not allow for sharing content with other BCs or BCFs. An example of a BCF in this class is Financial Metrics Lite (SolutionMatrix Ltd. 2010).
- Modern BCFs: Support the development of the BC in financial terms and often include specific criteria and methods for one or a few domains, such as RFID in the supply chain (Ivantysynova 2008, p.155). The relevant ones then need to be selected from a template or taxonomy, which has been pre-defined by domain experts. It is mostly possible to add new criteria and methods, however those can not be reused in other BCs or BCFs; they would need to be entered/copied manually. These BCFs may be implemented as a stand-alone Microsoft Excel file, but some of them are web applications that ease collaborative BCD, such as the Value Lifecycle Manager (SAP AG 2011).

The problems with the modern BCFs are still that (1) they are limited to the domains for which criteria and methods have been pre-defined by experts, (2) the criteria and methods need to be assigned explicitly to templates or taxonomies to facilitate their reuse and (3) the criteria, methods and their assignment to templates and taxonomies need to be maintained by the experts to enable the applicability of the BCF to new and changing domains.

	Traditional BCFs	Modern BCFs
Example	Financial Metrics Lite	Value Lifecycle Manager
Financial criteria	Included	Included
Domain-specific criteria/methods	Not included	Included, mostly limited to one domain
Selection of criteria/methods	No or little choice	Possible, e.g., select from template or taxonomy
Possible to define own criteria/methods	Sometimes possible, but only in financial terms	Mostly possible
Possible to reuse criteria/methods from other BCs/BCFs	Only manually, reusable components are fixed in the BCF	Only manually, reusable components are fixed in the BCF
Effort to maintain pre-defined criteria/methods	Low (financial methods have not changed much over the years)	High (experts need to develop templates or taxonomies)

Table 1. Simplified comparison of traditional and modern BCFs.

4 Support for the Dynamic Reuse of Criteria

This paper proposes a new mechanism applicable in BCFs, namely *dynamic reuse of BC components*. It distinguishes itself from the *static reuse of BC components* in modern BCFs in that the reusable, domain-specific criteria and methods do not need to be pre-defined by experts, but can be reused from earlier BCs, possibly developed by other BC developers. It is most likely that this will take place in large organizations with many BCs and limited restrictions for sharing strategic data. But it is not unthinkable that BCs will even be shared across organizational boundaries, possibly in an anonymized or aggregated form. To select criteria and methods (semi-)automatically from earlier BCs to be reused, a mechanism is needed, such as collaborative filtering (Hussein & Ziegler 2011). Collaborative filtering is common in other domains, such as online stores. For example, when looking at a certain book on Amazon, other books are recommended that were bought by other customers in combination with the book of interest. This principle can be applied to BCD as well. The development of a new BC is supported by means of an algorithm that scores criteria in similar earlier BCs. The algorithm globally works as follows (also see Figure 1):

1. It identifies the criteria which have been entered into the new BC.
2. It searches the database of earlier BCs, looking for BCs in which some of these criteria appear as well. These will be called the ‘matching BCs’. The criteria in these BCs need to be identifiable as such and should be structured hierarchically.
3. Each matching BC contains what will be called ‘potentially related criteria’. The strength of this potential relation is scored as follows. Criteria that are closer in the hierarchy to the place where the new and the earlier BC matched get more points.
4. For each of the potentially related criteria the scores are summed up over all the matching BCs. A limited set of the highest scoring criteria can then be recommended to the BC developer.

The scoring algorithm works as follows:

- siblings get 4 points (a sibling is on the same level *and* in the same branch of the hierarchy)
- parents and children get 3 points
- grandparents and grandchildren get 2 points
- all other criteria get 1 point

When a criterion occurs multiple times in the new BC (it is possible to repeat criteria in the criteria hierarchy), all matching BCs are scored only once with respect to that criterion. However, when an earlier BC matches to multiple different criteria in the new BC, all criteria in that earlier BC are scored for each matching criterion. Finally, when a matching criterion occurs multiple times in the earlier BC, that criterion will get scored multiple times, depending on its different positions in the hierarchy. For example, if ‘Turnover time’ would also occur below ‘Benefit ob3’ in Figure 1, it would receive 4+1=5 points.

New business case: RFID

Criteria
Impacts
Benefit nb1
Benefit nb2
Warehouse search time
Stock levels
Cost nc1
Contingencies
Risk nr1
Risk nr2

Earlier business case: barcode

Points	Criteria
1	Benefits
1	Benefit ob1
1	Benefit ob2
1	Benefit ob3
2	Costs
3	Warehousing costs
match	Stock levels
4	Turnover time
1	Cost oc1
1	Risks
1	Risk or1
1	Risk or2

Figure 1. Scoring criteria in the earlier BC based on matching criteria in the new BC

Compared to *static reuse of BC components*, the advantages of *dynamic reuse of BC components* are that it is not limited to a certain domain, the dimensions for reusability do not need to be defined explicitly, i.e., no formalization of the domain is needed. Moreover, there is little or no need for maintenance, the recommendations are always up-to-date and the recommendations may get better the more the system is used. This approach however also has some disadvantages: the quality of the recommendations may sometimes be insufficient and the recommender algorithm needs a starting point, i.e., it can only start looking for similar BCs as soon as some criteria have been entered in the new BC.

An important limitation of the algorithm used for the usability experiment presented in this paper, is that it only identifies similar BCs when they include criteria that are literally similar to the criteria in the new BC. This was not a problem during the usability experiment, because a rather homogeneous set of earlier BCs was used (see section 5). However, in the real world, BCs may be more heterogeneous, because they will include typos and other nuances that are common in natural language. Therefore, future versions of the algorithm should include natural language processing to enable more intelligent matching of new BCs to earlier BCs.

5 Method

5.1 Research Model

A usability experiment was set-up to compare different types of possible support for the reuse of criteria. Figure 2 shows the research model, which is based on the Technology Acceptance Model (F. D. Davis 1989; Venkatesh et al. 2003). It consists of the following dependent variables:

- Actual Efficiency: The time or effort the BC developer needs to develop a BC.
- Actual Effectiveness: The quality of the resulting BC, e.g., in terms of its reliability or acceptance by decision makers.
- Perceived Ease of Use: How easy to use the BCF is according to the BC developer.
- Perceived Usefulness: How useful the BCF is according to the BC developer.
- Intention to Use: How likely it is that the BC developer will use the BCF in his/her daily work (assuming that the BCF would be available and use would be voluntarily).

And the following independent variables:

- BCD Task: The type of BC that needs to be developed, e.g., determined by its size, domain, or complexity.
- BCD Experience: The experience of the BC developer in executing tasks similar to the BCD Task. Experience may concern BCD in general, but also the specific domain, e.g., industry or business processes affected.
- BCD Support: The information and tools that are available to the BC developer for completing the BCD Task.

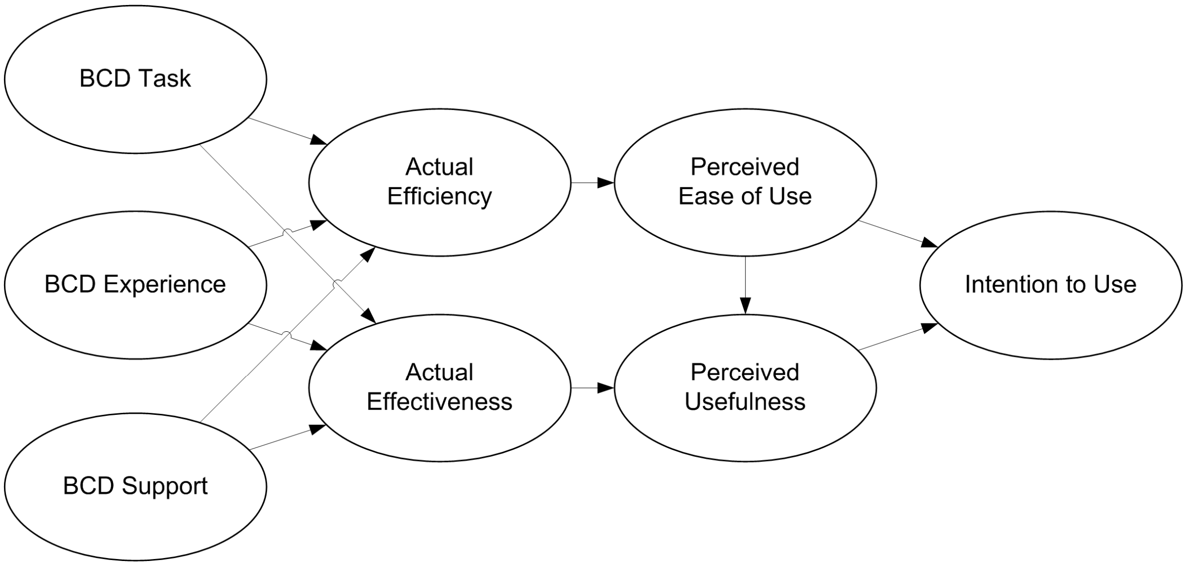

Figure 2. The research model.

5.2 Experimental Design

The usability experiment consisted of the following main phases:

- Pre-test: A questionnaire to investigate the BCD Experience variable;
- Introduction: A text explaining what needed to be done during the task;
- Benefit Definition Task: A simplified BCD task to compare different types of support for the reuse of criteria;
- Post-test: A questionnaire to investigate the perceptions of the participant;

The task was to define five benefit criteria for a certain domain, in this case the use of Radio-Frequency Identification (RFID) in the automotive supply chain. Normally, a BC would also consist of cost and risk criteria, but to keep it simple, the task was limited to the identification of benefits. There were three experimental conditions (variations of the independent variable ‚BCD Support‘):

- Recommendations (Appendix: Figure 4 to 8): Five benefits were recommended for reuse, based on the collaborative filtering principle presented in section 4. Each time a participant included a recommended benefit or came up with a benefit him/herself, the list of recommendations was updated automatically. This condition represents dynamic reuse of BC components.
- Templates (Appendix: Figure 9 to Figure 12): All benefits in the benefit database were listed, structured along the business processes to which they are related. Participants could enter benefits manually or include them from the templates. This condition represents static reuse of BC components, as common in modern BCFs such as the Value Lifecycle Manager (see section 3).
- No support (control-condition) (Appendix: Figure 13): All benefits needed to be invented and entered manually. This condition represents traditional BCFs, such as Financial Metrics Lite (see section 3).

In all conditions, one exemplary benefit criterion was provided, to help the participant get started. For the recommender system it was necessary to have at least one benefit criterion entered in the new BC, so that it could directly start identifying related earlier BCs.

The database of benefit criteria was developed through action research with domain experts in the RFID-based Automotive Network research project (Project RAN 2011; Reinhart et al. 2011). It consists of 330 benefits (criteria) for the use of RFID in the automotive supply chain. The benefits are structured hierarchically and divided over 12 BCs. Each BC represents the use of RFID to improve a certain business process in the automotive supply chain, e.g., ‘Goods Receipt’, ‘Transportation’ and ‘Storage’. In the recommendations-condition, benefits are recommended from this database. In the templates-condition, all benefits are drawn from this database and presented as hierarchically structured lists. Using such an expert validated dataset allowed for the reliable evaluation of the Actual Effectiveness. Moreover, by using it for both of the conditions where support was provided, the results became comparable.

A between subjects design was applied; each new participant was assigned automatically to the next experimental condition. Because not every participant completed the experiment, in the end the three groups did not have exactly the same size.

The *independent* variable ‘BCD Experience’ was measured with several questions during the pre-test (Appendix: Figure 1, Figure 2). The *dependent* variables were measured as follows:

- Actual Efficiency: The time in seconds was automatically measured from the start of the task until the participant continued to the next phase of the experiment.
- Actual Effectiveness: Benefits entered by participants were compared to the benefit database. For each exact match to the right category of Transportation benefits one point was assigned. For each exact match to another category half a point was assigned. For each benefit that was highly/somewhat similar (judged manually) to one in the Transportation category: one or half a point. All other benefits received zero points. The overall Actual Effectiveness for the participant was calculated as the mean of these points.
- Perceived Ease of Use: Some questions after the task had been completed (Appendix: Figure 14).
- Perceived Usefulness: Some questions after the task had been completed (Appendix: Figure 14).
- Intention to Use: Some questions after the task had been completed (Appendix: Figure 14).

The main hypotheses were that participants in the recommendations-condition would be the most efficient and those in the control-condition would be the least efficient. Participants in the templates-condition would be slightly more effective than those in the recommendations-condition, but those in the recommendations-condition would score higher on ease of use, usefulness and intention to use. For a more comprehensive list of the hypotheses please refer to Table 5 in section 6.2.

5.3 Pilot

For the execution of the usability experiment, a custom web application was developed (screenshots in Appendix: Figure 1 to Figure 14), to be used by the participants from their office desks or wherever and whenever they would want to participate. A pilot was done with two usability experts and two people who were representative for the target group of participants (see section 5.4). They were observed to see where they experienced trouble and after completing the experiment their experiences were discussed. Based on the pilot, several changes were made. The main change was the reduction of the number of criteria from 10 to 5, making the task easier for participants and reducing the overall time needed.

5.4 Sampling

Participants were invited among three types of people, who are representative for the target users of BCFs:

- Value Engineers: People who develop BCs for customers to support the sales of e.g., software solutions. These people are highly experienced in BC development and execution.
- Business Developers: People who develop BCs for improving the performance of the company internally, or its relations to partners. Although BC development is not necessarily part of their daily work, there is a high likelihood that there are many BCD experts among this group.
- Others: All other people who may have to develop a BC one day, e.g., to get an idea across to senior management. Most of these people have very little experience developing BCs.

The experiment was carried out within SAP, a global business software company. Invitations were sent by e-mail. Value Engineers were invited when they had the keyword ‚value engineer(ing)‘ in their profiles in the corporate address book. Business Developers were invited when they had the keyword ‚business develop(er/ment)‘ in their profiles. For the ‚Others‘ category, all employees of the research department were invited.

6 Results

6.1 Demographics

The participants were mainly male (77%) and highly educated (20% Bachelor, 53% Master, 17% PhD, or equivalent). The mean age was 33. Most participants usually work in Germany (46%), but several other countries were also well represented: France (11%), USA (8%), Great Britain (6%). 25% of the participants used the German language version of the experimental system; all others used the English version. The time was measured for every phase of the experiment. The phase requiring most time was the benefit definition task itself, with a median time of 2min 57sec. Table 2 shows the three types of participants, how many were invited, how many started with the experiment and how many completed it.

	Invited	Started	Completed	Completion Rate (Completed/Started)	Response Rate (Completed/Invited)
Value Engineer	115	52	29	56%	25%
Business Developer	335	90	49	54%	15%
Others	710	228	130	57%	18%
TOTAL	1160	370	208	56%	18%

Table 2. Different types of participants and their completion rates.

6.2 Hypotheses Testing

The recommendations-, templates- and control-conditions were compared on the five dependent variables. An analysis of variance (ANOVA) was conducted, with post-hoc tests using the Tukey correction (Table 3). In several cases, Levene's test of equality of error variances was significant and Welch's F is reported (Field 2005, p.350). The results show that people using the recommendations are more efficient (lower time value) than those using templates (hypothesis H1a, Table 3) or those without support (hypothesis H1b). People using recommendations or templates are more effective than those without support (H2b, H2c), but the templates are not more effective than the recommendations (H2a). The recommendations are perceived to be easier to use than the templates and the control-condition (H3a, H3b). The recommendations are more useful than the control-condition (H4b), however they are not perceived to be more useful than the templates (H4a). The intention to use the recommendations is higher than that for the templates (H5a) or the control-condition (H5b). To conclude, the recommendations are preferred over the other conditions, but do not differ from the templates in terms of effectiveness.

	N			M			SD			ANOVA	Pairwise Comparisons		
	R	T	C	R	T	C	R	T	C	F, Sig.	R vs T	R vs C	T vs C
Efficiency	76	64	75	163	256	245	111	163	127	**$H1_0$** F*(2,133)=12, p<.001	**H1a** p<.001	**H1b** p<.01	**H1c** p=.89
Effectiveness	377	318	341	0.91	0.89	0.74	0.19	0.21	0.29	**$H2_0$** F*(2,657)=45, p<.001	**H2a** p=.39	**H2b** p<.001	**H2c** p<.001
Ease of Use	71	61	57	4.1	3.7	3.6	0.87	0.67	0.86	**$H3_0$** F(2,186)=6.6, p<.01	**H3a** p<.05	**H3b** p<.01	**H3c** p=.72
Usefulness	70	60	54	3.8	3.5	3.2	0.81	0.99	1.3	**$H4_0$** F*(2,109)=4.8, p<.05	**H4a** p=.26	**H4b** p<.01	**H4c** p=.25
Intention	72	61	61	3.7	3.2	2.9	0.96	1.0	1.3	**$H5_0$** F*(2,122)=9.0, p<.001	**H5a** p<.05	**H5b** p<.001	**H5c** p=.46

Table 3. Analysis of variance. R=Recommendations, T=Templates, C=Control-condition, F=Welch's F. Hypotheses are indicated as Hx_y.*

In addition to the ANOVAs, bivariate correlations were computed between the independent variable 'Experience' and the dependent variables, as well as among the dependent variables (Table 4). The results show that more experience leads to a higher effectiveness (H6b), ease (H6c), usefulness (H6d), intention (H6e), but not to more efficiency, i.e., a lower time value (H6a). Less efficient people (higher time value) actually perform better (higher effectiveness) (H7a). A higher effectiveness leads to a higher ease of use (H8a), usefulness (H8b) and usage intention (H8c). A higher ease of use leads to a higher usefulness (H9a) and usage intention (H9b). By itself, a higher usefulness also leads to a higher usage intention (H10). These findings are conform the relations in the Technology Acceptance Model (F. D. Davis 1989; Venkatesh et al. 2003). However, it needs to be noted that only the last three correlations that were described are strong (r>.40).

	Efficiency	Effectiveness	Ease of Use	Usefulness	Intention to Use
Experience	**H6a** r(215)=.089, p=.19	**H6b** r(1036)=.07, p<.05	**H6c** r(189)=.16, p<.05	**H6d** r(184)=.24, p<.001	**H6e** r(194)=.32, p<.001
Efficiency		**H7a** r(1036)=-.11, p<.001	**H7b** r(189)=-.02, p=.77	**H7c** r(184)=.14, p=.064	**H7d** r(194)=.082, p=.27
Effectiveness			**H8a** r(921)=.08, p<.05	**H8b** r(902)=.085, p<.05	**H8c** r(947)=.17, p<.001
Ease of Use				**H9a** r(179)=.47, p<.001	**H9b** r(187)=.45, p<.001
Usefulness					**H10** r(182)=.66, p<.001

Table 4. Bivariate correlations for Task 1 (Reuse of Criteria). Hypotheses are indicated as Hx_y.

Table 5 shows all hypotheses and whether they can be accepted or should be rejected. The hypotheses for Actual Efficiency should be read carefully, because a higher Actual Efficiency means more time, which is worse. Thus, when there is a positive correlation between Actual Efficiency and e.g. Actual Effectiveness, this means that people spending more time are more effective.

$H1_0$	There is no difference in *Actual Efficiency* among the experimental conditions.	reject***
H1a	The *Actual Efficiency* of Recommendations is better than Templates.	accept***
H1b	The *Actual Efficiency* of Recommendations is better than Control group.	accept**
H1c	The *Actual Efficiency* of Templates is better than Control group.	reject
$H2_0$	There is no difference in *Actual Effectiveness* among the experimental conditions.	reject***
H2a	The *Actual Effectiveness* of Recommendations is lower than Templates.	reject
H2b	The *Actual Effectiveness* of Recommendations is higher than Control-condition.	accept***

H2c	The *Actual Effectiveness* of Templates is higher than Control-condition.	accept***
$H3_0$	There is no difference in *Perceived Ease of Use* among the experimental conditions.	reject**
H3a	The *Perceived Ease of Use* of Recommendations is higher than Templates.	accept*
H3b	The *Perceived Ease of Use* of Recommendations is higher than Control-condition.	accept**
H3c	The *Perceived Ease of Use* of Templates is higher than Control-condition.	reject
$H4_0$	There is no difference in *Perceived Usefulness* among the experimental conditions.	reject*
H4a	The *Perceived Usefulness* of Recommendations is higher than Templates.	reject
H4b	The *Perceived Usefulness* of Recommendations is higher than Control-condition.	accept**
H4c	The *Perceived Usefulness* of Templates is higher than Control-condition.	reject
$H5_0$	There is no difference in *Intention to Use* among the experimental conditions.	reject***
H5a	The *Intention to Use* of Recommendations is higher than Templates.	accept*
H5b	The *Intention to Use* of Recommendations is higher than Control-condition.	accept***
H5c	The *Intention to Use* of Templates is higher than Control-condition.	reject
H6a	*BCD Expertise* is positively correlated with *Actual Efficiency*.	reject
H6b	*BCD Expertise* is positively correlated with *Actual Effectiveness*.	accept*
H6c	*BCD Expertise* is positively correlated with *Perceived Ease of Use*.	accept*
H6d	*BCD Expertise* is positively correlated with *Perceived Usefulness*.	accept***
H6e	*BCD Expertise* is positively correlated with *Intention to Use*.	accept***
H7a	*Actual Efficiency* is positively correlated with *Actual Effectiveness*.	accept***
H7b	*Actual Efficiency* is positively correlated with *Perceived Ease of Use*.	reject
H7c	*Actual Efficiency* is positively correlated with *Perceived Usefulness*.	reject
H7d	*Actual Efficiency* is positively correlated with *Intention to Use*.	reject
H8a	*Actual Effectiveness* is positively correlated with *Perceived Ease of Use*.	accept*
H8b	*Actual Effectiveness* is positively correlated with *Perceived Usefulness*.	accept*
H8c	*Actual Effectiveness* is positively correlated with *Intention to Use*.	accept***
H9a	*Perceived Ease of Use* is positively correlated with *Perceived Usefulness*.	accept***
H9b	*Perceived Ease of Use* is positively correlated with *Intention to Use*.	accept***
H10	*Perceived Usefulness* is positively correlated with *Intention to Use*.	accept***

*Table 5. The hypotheses and results (***p<.001; **p<.01, *p<.05).*

6.3 Qualitative Results

The participants had the possibility to enter comments in an open text field, after each set of questions throughout the experiment. Those comments were clustered manually and some of the most noteworthy ones will now be presented.

With respect to *Usefulness*, the main comment (mentioned 18 times) was that the recommendations and templates reduce the participant's creativity. In the recommendations-condition, only 83% of the benefits entered were reused, while in the templates-condition 88% of the benefits were reused. This may imply that the recommendations-condition allows for more creativity than the templates-condition. Several participants (13x) who were assigned to the control-condition (no support) expressed some confusion when having to answer the ease of use, usefulness and intention to use questions for a system that did not provide any support at all. Four participants in the templates-condition mentioned that they found the list of benefits too long. With respect to *Ease of Use*, there were similar comments. Additionally, three participants mentioned that they found it confusing that the list of recommendations was updated after each benefit selection. Two participants in the templates-condition found it confusing that the list also contained non-relevant benefits. With respect to *Intention to Use*, six participants found it hard to answer the question, because they did not know what to compare the system to.

7 Conclusion

Business case development (BCD) is a complex activity, which can potentially be improved by supporting the reuse of investment criteria and valuation methods. The goal of this research was to improve the usefulness and usability of business case frameworks (BCFs) (Research Question 1), while limiting the effort required to develop and maintain static databases of reusable components. Therefore, an approach was proposed for the *dynamic reuse of BC components* and contrasted with *static reuse of BC components.* In the dynamic approach, the reusable, domain-specific criteria and methods do not need to be pre-defined by experts in templates and taxonomies, but can be reused from earlier BCs. To test whether support for dynamic reuse improves the usefulness and usability of BCFs (Research Question 2), a usability experiment was set up. Three types of support for the reuse of criteria were compared: (1) recommendations, based on collaborative filtering and representative for the dynamic approach, (2) templates, representative for the static approach common in the class of modern BCFs and (3) no support, representative for the class of traditional BCFs. For each type of support, the 'Actual Effectiveness', 'Actual Efficiency', 'Perceived Ease of Use', 'Perceived Usefulness' and 'Perceived Intention to Use' were measured. The task represented a simplified BCD activity, i.e., to define five benefit criteria for the use of RFID in the automotive supply chain, and was completed by 208 people. The results show that although the recommendations are as effective as the templates, they are preferred over the other conditions. Thus, to make BCD more convenient and to save costs in developing databases with reusable BC components, we recommend to apply the recommendations, or another type of support for dynamic reuse, in future research and future BCFs. The dynamic approach does not need to replace pre-defined templates and taxonomies, but could also be used as an extension to the static approach, to facilitate reuse in those cases where the new BC does not match a domain that has been described by experts.

Some limitations of this work are: (1) that all participants of the experiment work for the same company. However, the participants represented ages 20 to 60, work in 24 different countries and range from complete novices to BCD experts. (2) As the different types of support for reuse were integrated in an experimental system, the support was not perceived through the use of the original BCFs that were represented by the experimental conditions. Therefore, when answering the questions in the post-test, it may have been difficult for the participants to distinguish between the functionality being tested, and the experimental system. The integration of the BCFs' functionality in one experimental system was unavoidable to minimize the time needed per participant and to minimize the unintended effect of other variables. (3) The scoring algorithm used for the recommendations-condition, could only identify similar BCs when they included criteria that were literally similar to the criteria in the new BC. In future work, natural language processing techniques should be applied to increase the capability of the algorithm to deal with the heterogeneity of real world BCs.

In spite of these and some other limitations, we believe that this paper makes several significant contributions, because it (1) focuses on BCD as a human activity, with the related human-computer interaction issues, which is new, (2) presents a new approach to facilitate dynamic reuse, and (3) evaluates the approach in a large scale, empirical study, which is uncommon in the field of IS investment evaluation. Some good exceptions are (Ballantine & Stray 1998; Hochstrasser 1990; Serafeimidis & Smithson 2000; Ward et al. 2008; Willcocks & Lester 1991).

In future work, the dynamic approach should be integrated in a more comprehensive BCF, after which the BCF should be tested with BC developers, working on more comprehensive BCD tasks. One effect that should receive particular attention is the possibility of loss of creativity due to the recommendations. With respect to that, questions are to what extent 'serendipity', i.e., the possibility to use unexpected criteria, may compensate for a loss of creativity and to what extent the loss of creativity may be worth the improvement in terms of efficiency. Additionally, the dynamic approach may be applied to support the reuse of methods. For each criterion in the new BC, the algorithm could select methods that have been applied most often to valuate the respective criterion in earlier BCs.

References

Appendix: http://warhol.wiwi.hu-berlin.de/aigaion2/index.php/attachments/single/55

Bacon, J., 1992. The Use of Decision Criteria in Selecting Information Systems/Technology Investments. *MIS Quarterly*, 16(3), pp.335-353.

Ballantine, J. & Stray, S., 1998. Financial Appraisal and the IS/IT Investment Decision Making Process. *Journal of Information Technology*, 13(1), pp.3-14.

Ballantine, J. & Stray, S., 1999. Information Systems and Other Capital Investments: Evaluation Practices Compared. *Logistics Information Management*, 12(1/2), pp.78-93.

Davis, F.D., 1989. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3), pp.319-339.

Field, A., 2005. *Discovering Statistics Using SPSS*, London, UK: SAGE Publications Inc.

Gambles, I., 2009. *Making the Business Case - Proposals that Succeed for Projects that Work*, Farnham: Gower.

Hochstrasser, B., 1990. Evaluating IT Investments – Matching Techniques to Projects. *Journal of Information Technology*, 5, pp.215-221.

Hussein, T. & Ziegler, J., 2011. Situationsgerechtes Recommending. *Informatik-Spektrum*, 34(2), pp.143-152.

Irani, Z., 2002. Information Systems Evaluation: Navigating Through the Problem Domain. *Information & Management*, 40(1), pp.11 - 24.

Irani, Z., Ghoneim, A. & Love, P.E.D., 2006. Evaluating Cost Taxonomies for Information Systems Management. *European Journal of Operational Research*, 173(3), pp.1103 - 1122.

Ivantysynova, L., 2008. *RFID in Manufacturing: Mapping the Shop Floor to IT-Enabled Business Processes*. Berlin: Humboldt-Universität zu Berlin.

King, J.L. & Schrems, E.L., 1978. Cost-Benefit Analysis in Information Systems Development and Operation. *ACM Computing Surveys (CSUR)*, 10(1), pp.19-34.

Project RAN, 2011. RFID-based Automotive Network (RAN). Available at: www.autoran.de.

Van Putten, B.-J., Brecht, F. & Günther, O., 2011. Challenges in Business Case Development and Requirements for Business Case Frameworks. In 5th European Conference on Information Management and Evaluation (ECIME). Como, Italy.

Reinhart, G. et al., 2011. Wirtschaftlicher und ressourceneffizienter durch RFID? Methode für eine unternehmensübergreifende Bewertung der Wirtschaftlichkeit und der Ressourceneffizienz. *Zeitschrift für Wirtschaftlichen Fabrikbetrieb*, 106(4), pp.225-230.

Renkema, T.J.W. & Berghout, E.W., 1997. Methodologies for Information Systems Investment Evaluation at the Proposal Stage: a Comparative Review. *Information and Software Technology*, 39(1), pp.1 - 13.

SAP AG, 2011. Value Lifecycle Manager. Available at: https://sapvalueengineering.com/VLM2/client.page.

Schryen, G., 2010. Preserving Knowledge on IS Business Value. *Business & Information Systems Engineering*, (4), pp.233-244.

Serafeimidis, V. & Smithson, S., 2000. Information Systems Evaluation in Practice: A Case Study of Organizational Change. *Journal of Information Technology*, 15(2), pp.93-105.

SolutionMatrix Ltd., 2010. FREE Excel Financial Metrics Calculator. Available at: http://www.solutionmatrix.com/download-center.html.

Urbach, N., Smolnik, S. & Riempp, G., 2009. The State of Research on Information Systems Success - A Review of Existing Multidimensional Approaches. *Business & Information Systems Engineering*, 4, pp.315-325.

Venkatesh, V. et al., 2003. User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3).

Ward, J., Daniel, E. & Peppard, J., 2008. Building Better Business Cases for IT Investments. *MIS Quarterly Executive*, 7(1), pp.1-15.

Willcocks, L. & Lester, S., 1991. Information Systems Investments: Evaluation at the Feasibility Stage of Projects. *Technovation*, 11(5), pp.283-302.