

# Bias within systematic and non-systematic literature reviews: the case of the Balanced Scorecard

---

<b>Student:</b>	Steven Jansen
<b>Study:</b>	Master of Business Administration
<b>Track:</b>	Innovation and Entrepreneurship
<b>First Supervisor:</b>	T. De Schryver
<b>Second Supervisor:</b>	L.L.M. van der Wegen
<b>Date:</b>	26-09-2017

## Acknowledgements

At first I would like to sincerely thank Tom de Schryver for his incredible patience and guidance during this thesis project. Whenever I wrote parts of my thesis, he was always there with structural criticism and good feedback. Thanks to him, this thesis project became a success.

I would also like to sincerely thank my study advisor Charlotte Röring for the personal counselling and guidance, especially at the end of the thesis project. Charlotte Röring showed personal interest in my progress and offered guidance whenever I was having a difficult time.

At last I would like to thank my second supervisor Leo van der Wegen for his feedback and involvement in this thesis project.

Steven Jansen

## Abstract

According to Booth, Papaioannou & Sutton (2012) a literature review is important to any academic project as it is needed to fully understand the topic, by providing a bridge between the vast assortment of research available (Baumeister & Leary, 1997). When looking how literature reviews are performed a distinction can be made between two types of approaches: **Non-systematic** approaches and **systematic** approaches (Cronin, Ryan & Coughlan, 2008) (Okoli & Schabram, 2010). Authors like Tranfield et al. (2003) argue that **systematic** literature reviews are superior to **non-systematic** literature reviews, because **systematic** literature reviews include a stricter methodology when performing the reviewing process, yielding results that are supposedly less affected by **bias**. However since the literature on this discussion does not provide empirical evidence this research project delivered this by answering the following research question.

*“What is the impact of bias when selecting studies for a literature review as well as interpreting the results from these selected studies for non-systematic literature reviews as compared to systematic literature reviews?”*

To answer this question first a definition of what a **systematic** literature review is, is given. “A *systematic review attempts to collate all empirical evidence that fits pre-specified eligibility criteria in order to answer a specific research question. It uses explicit, systematic methods that are selected with a view to minimizing bias, [...]*” (Green, Higgins, Alderson, Clarke, Mulrow & Oxman, 2008, p. 13). A **non-systematic** literature review differs from a systematic literature review in that it is not obligated to be explicit about the methods that are used. It therefore is harder to identify sources of bias within non-systematic literature review which is further enhanced by the lack of formal quality appraisal.

To measure and compare the impact of bias on both **systematic** and **non-systematic** literature reviews a **non-systematic** case study was selected (Madsen, 2015). This case study was repeated using the same aim and boundaries, but with a **systematic** approach towards the reviewing process. The case study and the repeat review were then compared to each other to find differences in the process of selecting studies for inclusion in the review and interpreting, synthesizing and analysing the information of the included studies. Sources of **bias** and their impact were then identified for both the case study and the repeat review.

The results showed that both the **non-systematic** case study and the **systematic** repeat review were impacted by **bias** when selecting and interpreting studies (**place of publication bias, citation bias, etc.**) potentially affecting a significant number of the included studies in the repeat review and the case study. This evidence refutes claims made by authors like Tranfield, David & Palminder (2003) which deemed the **systematic** literature review to be superior. However, improving the **non-systematic** case study is hard, because it lacks a clear search strategy while the **systematic** repeat review does provide this. This thesis project showed a methodology for comparing **systematic** and **non-systematic** literature reviews, but since there was only **one comparison** made, more studies like this have to be performed to validate the methodology.

## Table of contents

Acknowledgements .....	I
Abstract .....	II
Table of figures.....	V
Table of tables .....	V
1 Introduction.....	1
1.1 Problem statement.....	1
1.2 Research goal .....	2
1.3 Research question .....	3
1.4 Sub-questions .....	3
1.5 Outline .....	4
2 Theoretical background.....	5
2.1 Typologies of literature reviews.....	5
2.2 Non-systematic literature reviews .....	7
2.2.1 Critical literature reviews .....	7
2.2.2 (Narrative) Literature review.....	8
2.2.3 Mapping literature reviews .....	9
2.2.4 Mixed methods literature reviews .....	10
2.2.5 Scoping literature reviews.....	11
2.2.6 State-of-the-art literature reviews .....	12
2.2.7 Overlapping or other types of non-systematic literature reviews.....	12
2.3 Systematic literature reviews.....	13
2.3.1 Systematic literature reviews.....	13
2.3.2 Meta-analysis literature reviews.....	16
2.3.3 Rapid Literature reviews .....	17
2.3.4 Umbrella literature review (review of reviews) .....	18
2.3.5 Overlapping or other types of systematic literature reviews .....	18
2.4 Differences between systematic and non-systematic literature reviews with regard to bias ...	19
2.4.1 A taxonomy of bias.....	19
2.4.2 Differences between systematic and non-systematic literature reviews.....	21
3 Methodology .....	23
3.1 Selecting and analysing a case study.....	23
3.1.1 Background on the subject of the case study .....	23
3.1.2 Finding a case study .....	24

3.2 Conducting the repeat review .....	26
3.3 Comparing the case study to the repeated study .....	29
4 Results .....	31
4.1 The case study .....	31
4.1.1 Introducing the case study .....	31
4.1.2 Description of the selection process of the case study .....	32
4.1.3 Interpretation of the information within included studies of the case study .....	36
4.2 The repeat review .....	37
4.2.1 The need for a systematic literature review .....	37
4.2.2 Defining a scope and research question accompanied by boundaries .....	37
4.2.3 Developing a review protocol with in-/exclusion criteria .....	38
4.2.4 Identifying and selecting research .....	39
4.2.5 Assessing the quality of the studies found .....	40
4.2.6 Extracting and synthesizing the data .....	41
4.2.7 Analysing, reporting and discussing the findings .....	44
4.3 Comparison of the case study and the repeat review .....	49
4.3.1 Comparison of both selection processes and outcomes .....	49
4.3.2 Comparison of both interpretation processes and outcomes .....	52
5 Conclusions, discussion, shortcomings and recommendations .....	54
5.1 Conclusions and discussion .....	55
5.1.1 Answering the sub-questions .....	55
5.1.2 Answering the main research question .....	57
5.1.3 Theoretical and practical relevance .....	57
5.2 Shortcomings .....	58
5.3 Recommendations .....	59
Appendices .....	60
Appendix A: Extraction form .....	60
Appendix B: Impact factor analysis .....	61
Appendix C: Data sheet of 11 studies included in both reviews .....	62
References .....	63
References of the repeat review .....	66

## Table of figures

Figure 1: Literature review continuum (Massaro et al., 2016) .....	2
Figure 2: The Balanced Scorecard (BSC) extracted from Kaplan & Norton (1992) .....	24
Figure 3: Number of articles concerning the Balanced Scorecard (Source: Scopus) .....	25
Figure 4: Year of publication of the included studies in the case study.....	33
Figure 5: type of data collection of the included studies in the case study.....	33
Figure 6: Hierarchy of evidence (Tranfield et al., 2003) .....	34
Figure 7: Network visualisation of the case study.....	35
Figure 8: In-/Exclusion criteria.....	40
Figure 9: Year of publication of the included studies in the repeat review .....	41
Figure 10: type of data collection of the included studies in the repeat review .....	42
Figure 11: Network visualisation of the repeat review .....	43
Figure 12: Studies included in both the case study and the repeat review .....	49
Figure 13: comparison on the year of publication of both reviews .....	50
Figure 14: comparison of the impact factor analyses .....	51

## Table of tables

Table 1: Typologies of literature reviews .....	6
Table 2: SALSA elements of a critical literature review.....	8
Table 3: SALSA elements of a (narrative) literature review .....	8
Table 4: SALSA elements of a mapping literature review .....	9
Table 5: SALSA elements of a mixed methods literature review .....	10
Table 6: SALSA elements of a scoping literature review .....	11
Table 7: SALSA elements of a state-of-the-art literature review .....	12
Table 8: Similarities between approaches towards systematic literature reviewing .....	14
Table 9: Similarities between approaches towards systematic literature reviewing .....	15
Table 10: SALSA elements of a meta-analysis literature review .....	16
Table 11: SALSA elements of a rapid literature review .....	17
Table 12: SALSA elements of a umbrella literature review.....	18
Table 13: A taxonomy of bias .....	21
Table 14: Free-line-coding (example).....	29
Table 15: Authors with multiple contributions to the case study.....	32
Table 16: Cluster description of the case study .....	35
Table 17: PICOC criteria.....	38
Table 18: Search terms (Boolean Logic) .....	39
Table 19: Cluster description of the repeat review.....	43
Table 20: Coding scheme of the repeat review.....	45
Table 21: Analysis of the studies included in both reviews.....	49
Table 22: study designs of included studies.....	50
Table 23: Comparison of the found clusters in both reviews .....	52
Table 24: Impact factor analysis of the repeat review .....	61
Table 25: Impact factor analysis of the case study (Madsen, 2015) .....	61

# 1 Introduction

This chapter describes the subject of this master thesis. First the current situation is described in the problem statement followed by the research goal and the research questions. At the end of this chapter an outline of this thesis is also described.

## 1.1 Problem statement

According to Booth et al. (2012) a literature review is important to any academic project as it is needed to fully understand the topic and thus providing a background (Okoli & Schabram, 2010). Webster & Watson (2002) are going even further by claiming it is an essential part to every academic project. Cronin et al. (2008) further emphasizes on this subject by stating that the goal of literature review is to bring the reader up-to-date with current literature on a topic. This is done by providing a bridge between the vast assortment of research and the reader (Baumeister & Leary, 1997). Furthermore Cronin et al. (2008) and Webster & Watson (2002) both state that literature reviews can form the basis for future research areas. While Rowe (2014) also states that literature reviews can be used to critically examine past research in relation to the topic.

But, what really is a literature review? Hart (1998, p. 1) defines a literature review as followed: “A literature review is an objective, thorough summary and critical analysis of the relevant available research and non-research literature on the topic being studied.”<sup>1</sup> Fink (2005) defined a literature review as: “A systematic, explicit, and reproducible method for identifying, evaluating and synthesizing the existing body of completed and recorded work produced by researchers, scholars and practitioners.”<sup>2</sup> Torraco (2005) has given the following definition of a literature review “A form of research that reviews, critiques, and synthesizes representative literature on a topic in an integrated way such that new frameworks and perspectives on the topic are generated.”<sup>3</sup> Beside these given definitions there are many more but all share some topics which are: identifying relevant literature, reviewing/evaluating the literature, and synthesizing/summarizing the literature. This is supported by Arksey & O’Malley (2005) that states: “This rapid growth in undertaking reviews of the literature has resulted in a plethora of terminology to describe approaches that, despite their different names, share certain essential characteristics, namely, collecting, evaluating and presenting the available research evidence.”

When looking how literature reviews are performed a distinction can be made between two types of approaches: **Non-systematic** (narrative/traditional) approaches and **systematic** approaches (Cronin et al., 2008) (Okoli & Schabram, 2010). While Fink (2005) uses the word “systematic” in its definition for literature reviews in general, Booth et al. (2012) comments on this by claiming that even non-systematic literature reviews have a certain degree of systematics embedded into them. Massaro, Dumay & Guthrie (2016) explains this degree of systematics by introducing the “literature review continuum” (see figure 1) where non-systematic approaches have less rules to follow than systematic approaches when performing the reviewing process.

---

<sup>1</sup> Hart (1998) extracted from Cronin et al. (2008)

<sup>2</sup> Fink (2005) extracted from Booth et al. (2012)

<sup>3</sup> Torraco (2005) extracted from Rocco & Plakhotnik (2009)

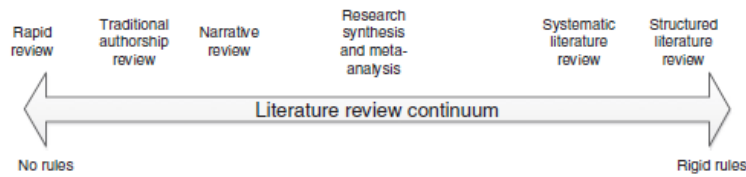


Figure 1: Literature review continuum (Massaro et al., 2016)

Mulrow (1994) argues that performing a literature review in a systematic manner is a search for the whole truth instead of just a part of it (bias), making it a fundamentally scientific activity. Tranfield et al. (2003) state that: “systematic literature reviews differ from non-systematic literature reviews because they adopt a replicable scientific and transparent process which is aimed at minimizing bias. This is achieved by performing exhaustive literature searches of published and unpublished studies and by clearly stating the reviewer’s decisions, which procedures were being used and the conclusions” (Tranfield et al., 2003). Tranfield et al. (2003) clearly argue for performing systematic literature reviews over non-systematic literature reviews by stating that a systematic literature review helps creating a reliable knowledge base. Scandura & Williams (2000, p. 1263) argues that: “without rigor, relevance in management research cannot be claimed”. Analysing and identifying previous research materials creates a prerequisite for creating knowledge, but with a more systematic approach it is possible to make an advance on the status quo and reduce subjectivity (Massaro, 2016). Petticrew & Roberts (2006) also deems non-systematic literature reviews to be inferior to systematic literature review by claiming that non-systematic literature reviews do not apply scientific principles when carrying out the reviewing process. This according to Petticrew & Roberts (2006) leads to an unrepresentative sample of studies included into the reviewing process which leads to an unrepresentative synthesis and conclusions. Petticrew & Roberts (2008) observed that non-systematic literature reviews were often performed by high profile researchers/expert which might be an indicator of biased and unreliably representations of the results from the performed literature review.

While the authors above argue that a systematic literature review is superior to a non-systematic one, there is no hard empirical evidence that the result of literature review is per definition biased when performed in a non-systematic way. While the arguments of authors like Tranfield et al. (2003), Petticrew & Roberts (2006) and Mulrow (1994) might be valid, both approaches are not tested side-by-side on the same topic under the same conditions. Massaro et al. (2016) argues that even systematic literature reviews can be biased as the researcher chooses the body of research from which the literature review is being performed. Therefore it isn’t possible to write off non-systematic literature reviews as inferior, until sufficient evidence, in the form of side-by-side comparison, has been provided.

## 1.2 Research goal

As mentioned in the problem statement (section 1.1) there is not sufficient evidence to claim that systematic literature reviews are superior to non-systematic literature reviews. Therefore it is necessary to assess the impact of bias when selecting studies to include in the literature review and the bias when interpreting the results from the selected studies, for both the systematic and non-systematic approach. Given this research gap the following research goal can be formulated

*“The goal of this master thesis project is to investigate the impact of bias when selecting studies for a literature review as well as interpreting the results from these selected studies for non-systematic literature reviews as compared to systematic literature reviews. Therefore a thorough analysis on both approaches will be conducted followed by a side-by-side comparison based on a case under the same conditions.”*

### 1.3 Research question

After developing the goal for the research project it is important to translate the goal towards a workable and answerable question. For this research project the research question is as follows:

*“What is the impact of bias when selecting studies for a literature review as well as interpreting the results from these selected studies for non-systematic literature reviews as compared to systematic literature reviews?”*

### 1.4 Sub-questions

To make the research question more workable sub-questions have been formulated which are more easy to answer. A short explanation is provided on how the answer to the question is achieved. Out of the research question the following sub-questions are formulated:

1. *“What types of non-systematic and systematic literature review methods are there?”*

The answer to this question is provided by searching for all the different types of literature reviews. Books on literature reviews provide a starting point of which by snowballing more types of literature reviews are found. After identifying the different types of literature reviews they are each explained in terms of how they work (how those literature reviews should be performed).

2. *“How does bias influence systematic and non-systematic literature reviews?”*

This question is answered by identifying strong and weak sides of the various types of systematic and non-systematic literature reviews which were identified earlier. The strong sides are analysed primarily by looking at parts of the reviewing process in which bias is minimized. The weak sides are analysed by looking at limitations for each type of literature review accompanied by a taxonomy of potential sources of bias affecting literature reviews.

3. *“What differences do occur when selecting studies for the literature review between a non-systematic literature review and a systematic literature review?”*

To answer this question a side-by-side comparison is performed. Therefore a case study was selected which is a non-systematic literature review. The subject of the case study as given by the supervisor is the **Balanced Scorecard** (BSC), which will be explained in section 3.1. This case study is thoroughly analysed on how the review was performed and under what conditions. After this the conditions of the case study are recreated and then a systematic literature review is carried out under the same conditions on the same topics. By analysing the differences in both approaches in terms of selecting research material for the review possible biases are identified. This analysis is aided by using network visualisations which couples authors to each other providing an overview of research

4. *“What differences do occur when interpreting the results from the selected studies material a non-systematic literature review and a systematic literature review?”*

This question will be answered by first using coding to find general themes in the selected research materials of the already existing non-systematic literature review. Secondly, the synthesis of these research materials by the authors will be compared to the themes found to find differences. After this a thematic analysis is performed on the recreated systematic literature review to find general themes within the selected research materials. At last the themes found of both the existing non-systematic literature review are compared to the recreated systematic literature review.

## 1.5 Outline

This report starts of by providing an introduction in which the main problem are stated followed by the research goal and (sub) question(s). After that the different types of literature reviews are identified in terms of strengths and weaknesses in terms of bias. Then the methodology is described starting with a case selection for the side-by-side comparison of both non-systematic and systematic literature reviews. Followed by a description of the reviewing process when using a systematic approach in contrast to the non-systematic case study. This also includes a description of the way in which network visualisations are used to help finding differences between both approaches. After conducting the reviewing process in a systematic fashion the results can be compared with the case study. The results of this comparison are first described in a descriptive manner, stating the origin of the research materials, dominant authors, etc. of both the case study and the systematic recreated case study. Then the results of the network visualisation are analysed and compared between both approaches to find differences in the selection of research materials. After that the results of the thematic analysis are compared between both approaches to find differences in the interpretation of the selected studies. At last conclusions are drawn and the research (sub-)question(s) are being answered as well as a discussion, stating the weaknesses of this report and its contribution towards the research field as well as making recommendations.

## 2 Theoretical background

This chapter is focused on identifying and analysing the various types of literature reviews both non-systematic and systematic. This is done in order to offer insight into how the types of literature review are being performed. The literature has been searched in a non-systematic fashion (because of limited time) via Scopus, Web of Science, Google Scholar and the library of the university of Twente. The search process was aimed at finding literature on the reviewing processes, a typology of literature reviews and biases associated with literature reviews. The found literature provided an image of the main concept within the area of literature reviewing, however it might be possible due to time constraints that certain concepts were omitted. By analysing these reviewing processes it becomes clear where there might be bias impacting the results of the literature review. Therefore an understanding and explanation of bias impacting literature reviews is also needed. This theoretical background offers a base on which the case study and the methodology for recreating the case study in a systematic fashion can be built.

### Chapter Summary

- *“A systematic review attempts to collate all empirical evidence that fits pre-specified eligibility criteria in order to answer a specific research question. It uses explicit, systematic methods that are selected with a view to minimizing bias, thus providing more reliable findings from which conclusions can be drawn and decisions made.”(Green et al., 2008, p. 13)(see subsection 2.3.1).*
- A non-systematic literature review differs from a systematic literature review in that it is not obligated to be explicit about the methods that are used. It therefore is harder to identify sources of bias within non-systematic literature review which is further enhanced by the lack of formal quality appraisal (see subsection 2.4.2)
- Publication and associated biases (Song et al., 2010) are influencing the selection process of studies for literature reviews. These biases refer to studies which are only published when its effects are significant and in the desired direction, thus omitting less significant studies (or in the wrong direction) creating an over-exaggerated image of the effect and a skewed literature review (see subsection 2.4.1).
- Interpretation bias refers to researchers (or reviewers) ability to synthesize, judge and weigh the results found in a study. Two researchers of different backgrounds might look at the same result in a different way thus drawing different conclusions based on their own background (MacCoun, 1998) (see subsection 2.4.1).
- Funding bias refers to when studies on, for example products, produce favourable outcomes for the company that makes the product because the company funds the study and its researchers (Booth et al., 2012) (Lexchin, Bero, Djulbegovic & Ottavio, 2003) (see subsection 2.4.1).

### 2.1 Typologies of literature reviews

When looking at figure 1 which presents the literature review continuum (Massaro, 2016) it offers a hidden definition for what can be considered as non-systematic literature reviews. Non-systematic literature reviews can be viewed as literature reviews which do not have to subject to as many rules

and rigour as systematic literature reviews. In figure 1 some examples are given of different types of literature reviews by Massaro (2016). Grant & Booth (2009), Malidou (2014), Paré, Trudel, Jaana & Kitsiou (2015) and Wickremasinghe, Kuruville, Mays & Avan (2015) further explored the different types of literature reviews and both came up with a typology list of literature reviews (see table 1).

<b>Grant &amp; Booth (2009)</b>	<b>Malidou (2014)</b>
Critical review	Critical/integrative review
Literature review	Meta-analysis
Mapping review	Rapid review
Meta-analysis	Realist review
Mixed methods review	Scoping review
Overview	Systematic review
Qualitative systematic review	Umbrella review
Rapid Review	
Scoping review	
State-of-the-art review	
Systematic review	
Systematic search and review	
Systematized review	
Umbrella review	
<b>Paré et al. (2015)</b>	<b>Wickremasinghe et al. (2015)</b>
Critical review	Annotated bibliography
Descriptive review	Evidence map
Meta-analysis	Evidence paper
Narrative review	Literature review
Qualitative systematic review	Mixed methods research synthesis
Realist review	Rapid review
Scoping/mapping review	Review of reviews
Theoretical review	Scoping review
Umbrella review	State-of-the-art review
	Systematic review

**Table 1: Typologies of literature reviews**

Between the typologies given by Grant & Booth (2009), Malidou (2014), Paré et al. (2015) and Wickremasinghe et al. (2015) there are quite a lot of similarities. Paré et al. (2015) is mainly focussed on research in information systems which might limit the ranges of different kinds of literature reviews, while Grant & Booth (2009) don't have this possible limitation. The same possible limitation might affect Malidou (2014) and Wickremasinghe et al. (2015) as they have been written for the healthcare sector. Nevertheless the similarities among different sector might indicate rather universal typologies of literature reviewing. Grant & Booth (2009), Malidou (2014), Paré et al. (2015) and Wickremasinghe et al. (2015) do provide descriptions of all the typologies mentioned in table 1. Grant & Booth (2009), Malidou (2014) and Wickremasinghe et al. (2015) are even going further by explicitly stating perceived strengths and weaknesses of the different kinds of literature review.

Wickremasinghe et al. (2015) also offers insight into why there are differences between the typologies given by the authors by also stating synonyms of the literature review types.

Wickremasinghe et al. (2015) states that a scoping review can also be known as a critical review which for example the other authors distinguish between. This indicates that there is not a 'golden standard' of literature review typologies. The typologies given by these authors will serve as a

starting point for describing the various non-systematic literature reviewing methods along with their strengths and weaknesses. The systematic types of literature reviewing methods will be described and analysed hereafter in section 2.3.

## 2.2 Non-systematic literature reviews

This section describes and analyses the typologies of literature reviews that are non-systematic. Since there is no single form of a non-systematic literature review, descriptions are made for the various typologies of table 1. For the analysis Booth et al. (2012) provided a framework with which different types of literature reviews can be compared. This framework is named SALSA which stands for: “Search (S), Appraisal (AL), Synthesis (S) and Analysis (A) which signify elements which are present within every type of literature review (Booth et al., 2012).

- **Search:** This element concerns the search for literature which according to Malidou (2014) includes a specification of the sources which are going to be used, the search strategy, time constraints and the possible inclusion of a scope on what to search for within the found literature.
- **Appraisal:** This element concerns the quality assessment of the found literature which according to Malidou (2014) includes a specification of criteria on which literature is being judged, whether or not a standard instrument is being used and how many of the found studies literature are being reviewed (sampling or total).
- **Synthesis:** This element concerns combining the summarized data found in the literature with the new insight from the writer to answer a pre-defined question (Malidou, 2014). According to Malidou (2014) this can be done in a narrative fashion, with a possible addition of a graphical diagram or a tabular synthesis.
- **Analysis:** This element concerns how the synthesized literature is being analysed and what the literature review seeks out to deliver in terms of outcome. Literature reviews might seek out to analyse the synthesized literature to create a conceptual model (Malidou, 2014) or might deliver recommendations based on a numerical analysis.

Grant & Booth (2009) and Malidou (2014) and Wickremasinghe et al. (2015) all used this method for developing their typologies. Therefore this framework is used to describe the various types of non-systematic literature reviews in this section.

### 2.2.1 Critical literature reviews

According to Grant & Booth (2009) a critical literature review is focused on extensively researching studies that goes beyond only a description of the studies. Rather a critical literature review often has a certain degree of analysis and even conceptual innovation (Grant & Booth, 2009). Paré et al. (2015) adds to this by stating that critical literature reviews are aimed at analysing the extant literature to reveal possible weaknesses, contradiction, controversies or inconsistencies. Rather than only comparing extant literature to each other, critical literature reviews classify studies towards chosen or developed criterion. Croom, Romano & Giannakis (2000) gives an example of these criterions by classifying studies found for their critical literature review by a content-criterion and a methodology-criterion.

SALSA elements	Grant & Booth (2009)	Malidou (2014)
Search	Seeks to identify the most significant	Significant literature is identified without

	studies	recording the search process
Appraisal	No formal quality assessment. Attempts to evaluate according to contribution	Evaluation of the literature is based on contribution, no criteria are specified
Synthesis	Typically narrative, perhaps conceptual or chronological	Narrative, written in a chronological or conceptual order
Analysis	Seeks to identify conceptual contribution to embody existing or derive new theory	Mainly identifies significant conceptual components or develops new theories

Table 2: SALSA elements of a critical literature review

### Strengths

- Critical literature review can identify problems, discrepancies or areas in which existing literature is not trustworthy and therefore might add towards development of that research area (Paré et al., 2015).
- A critical literature reviews value lies in the ability to focus on extensively reviewing the selected literature to really identify the value of its contribution (Grant & Booth, 2009).
- The outcome of a critical literature review might lead to a new phase of conceptual development of theories (Grant & Booth, 2009).

### Weaknesses

- There are no formal requirements of specifying inclusion or exclusion criteria during a quality assessment (Grant & Booth, 2009) (Malidou, 2014) (Paré et al., 2015).
- Due to the narrative approach in which the literature is synthesized the interpretation of the literature is prone to subjectivity, only conceptual or chronological criteria might be used (Grant & Booth, 2009) (Malidou, 2014).

### 2.2.2 (Narrative) Literature review

According to Paré et al. (2015) a narrative literature review simply attempts to find what has been written about a certain subject. This most often does not involve a comprehensive search strategy neither does it specify how the primary studies for the review were found and selected (Paré et al., 2015). According to Grant & Booth (2009) a narrative literature review focusses mainly on recent or current literature, which is readily available to the researchers (Wickremasinghe et al., 2015). The completeness of and comprehensiveness of the search and analysis elements may vary greatly and differs per narrative literature review (Grant & Booth, 2009).

SALSA elements	Grant & Booth (2009)	Wickremasinghe et al. (2015)
Search	May or may not include comprehensive searching	Seeks to identify significant papers
Appraisal	No formal quality assessment. Attempts to evaluate according to contribution	Limited quality assessment, most likely critical appraisal of contribution of the found literature
Synthesis	Typically narrative	Narrative
Analysis	Analysis may be performed to find themes or concepts	Seeks out to provide areas of consensus and debate within a research area

Table 3: SALSA elements of a (narrative) literature review

## Strengths

- The strength of a narrative literature review is that it offers a base for building on previous work, thus avoiding duplication, in a research area by offering a summation (Grant & Booth, 2009).
- A narrative literature review also helps to identify omissions and gaps within the research area (Grant & Booth, 2009).

## Weaknesses

- There is not an explicit intent to include as much as literature as possible in narrative literature reviews (Grant & Booth, 2009) (Paré et al., 2015) (Wickremasinghe et al., 2015).
- Authors may only select literature which supports their case or their preferred hypothesis (Grant & Booth, 2009).
- Replication of a narrative literature review is deemed to be impossible as information on how primary studies were searched and selected is not available (Dijkers, 2009)<sup>4</sup>.
- Due to the narrative approach in which the literature is synthesized the interpretation of the literature is prone to subjectivity (Grant & Booth, 2009).

### 2.2.3 Mapping literature reviews

According to Grant & Booth (2009) mapping literature reviews are aimed at categorizing existing literature on a certain topic. An example of this can be given by looking at the study of O’Cathain, Thomas, Drabble, Rudolph & Hewison (2013), as it came up with 5 categories in which the topics of the included literature were divided. Dicheva, Dichev, Agre & Angelova (2015) came up with 6 dimensions in which the topics of their literature review are categorized. In mapping the literature by categorizing, gaps in literature can be identified which indicates the need for further and/or primary research (Grant & Booth, 2009).

<b>SALSA elements</b>	<b>Grant &amp; Booth (2009)</b>	<b>Wickremasinghe et al. (2015)</b>
Search	Aimed at finding all literature, restricted by time or scope	A comprehensive search for literature, within a specific time frame
Appraisal	No formal quality assessment	Limited quality assessment
Synthesis	Graphical or tabular	Graphics and tables
Analysis	Characterizes literature perhaps by study design and other key-features. May identify a need for primary or secondary research	Provides an overview of key themes or results within a research area and identifies research gaps.

Table 4: SALSA elements of a mapping literature review

## Strengths

- Mapping literature reviews are able to identify narrower research question for more in-depth literature reviews (Grant & Booth, 2009) (Wickremasinghe et al., 2015).
- Mapping literature reviews also offer insight into which resources (time, researchers, etc.) are needed to undertake new more in-depth literature reviews (Grant & Booth, 2009).

<sup>4</sup> Dijkers (2009) extracted from Paré et al. (2015)

## Weaknesses

- Mapping literature reviews lack a comprehensive quality assessment which while it mostly categorizes literature by their study design and not on other elements (Grant & Booth, 2009).
- The analysis of the found literature is mostly on a broad descriptive level which might oversimplify the conclusions which are drawn and masks variation between the results of the individual studies (Grant & Booth, 2009).

### 2.2.4 Mixed methods literature reviews

According to Grant & Booth (2009) a mixed methods literature review combines several types of literature reviews where usually a systematic method is included. However, it has been classified as non-systematic since it is not obligated to include systematic methods. This type of literature review is aimed at bringing together quantitative data together with a qualitative literature review.

Wickremasinghe et al. (2015) states that a mixed methods literature review provides a comprehensive summary of the evidence within both quantitative and qualitative literature which underpins policy decisions of practitioners.

<b>SALSA elements</b>	<b>Grant &amp; Booth (2009)</b>	<b>Wickremasinghe et al. (2015)</b>
Search	Aimed at finding all study through a quantitative and/or qualitative search strategy or very thoroughly searching through different databases	Aimed at finding both quantitative and qualitative literature
Appraisal	Makes use of a generic research appraisal instrument or may have separate appraisal processes with corresponding checklists	Formal quality assessment is essential and uses documented appraisal processes
Synthesis	Usually narrative and/or tabular but may contain graphical elements to include quantitative literature	Narrative, graphical or tabular depending on the included literature (quantitative and/or qualitative)
Analysis	Compares literature with each other while looking for correlations between characteristics by means of a gap analysis	A full map is given of the found quantitative and qualitative literature to answer a research question.

Table 5: SALSA elements of a mixed methods literature review

## Strengths

- A mixed methods literature review capitalizes on the strengths of a systematic literature review by providing a documented form of appraisal of the literature (Grant & Booth, 2009).
- The mixed methods literature review might also provide a more complete picture of the whole body of research on a certain topic including multiple types of literature (both quantitative and qualitative) (Grant & Booth, 2009) (Wickremasinghe et al., 2015).

## Weaknesses

- While a mixed methods literature review seeks to combine the strengths of multiple types of literature reviewing it will also compound methodological challenges of bringing these together, since there doesn't exist any consensus on how to do this (Grant & Booth, 2009).

- Bringing together the literature found into a synthesis may also prove to be difficult as the quantitative studies might be structured very differently when compared to qualitative studies (Grant & Booth, 2009). Even if the different studies might ask the same research question they might also be conducted in different paradigms (Grant & Booth, 2009) thus making it hard to interpret and combine the results.
- A mixed methods literature review is also very time consuming and resource intensive (requires a large staff and many databases to acquire literature from) (Wickremasinghe et al., 2015).

### 2.2.5 Scoping literature reviews

According to Grant & Booth (2009) and Paré et al. (2015) a scoping literature review offers a preliminary assessment on potential size and scope of literature on a certain topic. Malidou (2014) adds to this by stating that scoping literature reviews also identify which types of evidence (empirical, etc.) is mostly available in the research area as well as what the main sources are.

<b>SALSA elements</b>	<b>Grant &amp; Booth (2009)</b>	<b>Malidou (2014)</b>	<b>Wickremasinghe et al. (2015)</b>
Search	Completeness of the search is determined by time/scope constraints.	A comprehensive or complete search defined by the scope of the review	Search undertaken on a certain topic (constrained by the scope and time)
Appraisal	No formal quality assessment	Informal quality assessment or no quality assessment	-
Synthesis	Tabular with narrative commentary	Tabular with narrative commentary	Narrative and tables
Analysis	Characterizes quantity and quality of literature by key features such as study design	Mapping quantity and quality of the literature for identifying research gaps	To summarize and disseminate findings in literature for identifying research gaps

Table 6: SALSA elements of a scoping literature review

#### Strengths

- Scoping literature reviews maps the research area on a certain topic (Grant & Booth, 2009) (Paré et al., 2015).
- Scoping literature reviews may determine the need for a full systematic literature review on a certain topic (Grant & Booth, 2009) (Paré et al., 2015).
- Scoping literature reviews may also offer insight by summarizing and disseminating the research finding across the studies in the research area of a certain topic (Arksey & O'Malley, 2005) (Wickremasinghe et al., 2015).

#### Weaknesses

- There is no formal quality assessment of the found literature (Grant & Booth, 2009) (Malidou (2014) (Wickremasinghe et al., 2015). Paré et al. (2015) does argue that at least two researchers should carry out an assessment of quality for excluding studies from the review which do not address toward the research question of the scoping literature review.

- Scoping literature reviews are limited by time constraints and lack the rigour of a systematic literature review to clarify all concessions being made because of this (Grant & Booth, 2009) (Wickremasinghe et al., 2015)

### 2.2.6 State-of-the-art literature reviews

According to Grant & Booth (2009) a state-of-the-art literature review is more focussed on current matters within a research area of a certain topic in contrast to other literature reviews which take a more retrospective approach. According to Grant & Booth (2009) state-of-the-art literature reviews in their analysis will mostly highlight future research topics. This is supported by Nudurupati, Bititci, Kumar & Chan (2010) which specifically states six new topics of research in the conclusion of their study.

<b>SALSA elements</b>	<b>Grant &amp; Booth (2009)</b>	<b>Wickremasinghe et al. (2015)</b>
Search	Aimed at a comprehensive search of current literature	Aimed at finding the most recent literature
Appraisal	No formal quality assessment	Limited quality assessment
Synthesis	Typically narrative but may be accompanied by tables	Narrative, graphical or tabular
Analysis	Analyses the current state of the research area on a topic and highlights future research topics	Analysis on the most recent literature of a topic to provide evidence to support policymakers

Table 7: SALSA elements of a state-of-the-art literature review

#### Strengths

- The strength of a state-of-the-art literature review lies in its ability to provide a current image of developments for those who are new to a research area, instead of having to read multiple recent articles to achieve knowledge of these developments (Grant & Booth, 2009).

#### Weaknesses

- A weak side of state-of-the-art literature reviews reveals itself when a topic of a research area is extensively studied in the past but not in recent times thus creating a skewed total image of the topic (Grant & Booth, 2009).
- With state-of-the-art literature reviews the synthesis and analysis of the literature depends on a specific time-horizon of the included review, which potentially understate findings of prior literature and overstates current popular topics of a research area (Grant & Booth, 2009).

### 2.2.7 Overlapping or other types of non-systematic literature reviews

#### Descriptive literature reviews

According to Paré et al. (2015) a descriptive literature review seeks out to find patterns among literature on a certain topic. In doing so and to assure generalizability of the results, a descriptive literature review collects a representative sample of studies from a research area of which they codify and analyse the numerical data (Paré et al., 2015). The numerical data found among the included literature consist of certain characteristic like: the publication year, research methods used, data collection techniques, etc. (Paré et al., 2015). Paré et al. (2015) argues that descriptive literature

reviews are treating included studies as units of analysis with which they try to give a representation of the state-of-art of the research area.

### Overview

According to Grant & Booth (2009) an overview is a generic term which is used for any type of summary of literature that attempts to describe its characteristics. It might offer newcomers to a topic some insight but lacks any kind of systematic methodology or reporting of decisions made (Grant & Booth, 2009).

### Theoretical literature reviews

According to Paré et al. (2015) a theoretical literature review seeks to combine the conceptual contribution of empirical studies with existing theoretical studies. This type of literature reviews does share the SALSA characteristics with a critical literature review (see sub-section 2.2.1) except for the analysis. The difference in the analysis is that the goal of a theoretical literature review is to provide a theoretical foundation for a topic in a research area which lacks one (Paré et al., 2015).

## 2.3 Systematic literature reviews

This section describes and analyses the typologies of literature reviews that are systematic. The description provided makes use of the types of literature reviews mentioned in section 2.1. The SALSA elements, as mentioned in section 2.2 will be used to describe the types of systematic literature reviews.

### 2.3.1 Systematic literature reviews

When looking at systematic literature reviews Green et al. (2008) developed a definition and description which sets it apart from the non-systematic ones. Green et al. (2008) defines a systematic literature reviews as: *“A systematic review attempts to collate all empirical evidence that fits pre-specified eligibility criteria in order to answer a specific research question. It uses explicit, systematic methods that are selected with a view to minimizing bias, thus providing more reliable findings from which conclusions can be drawn and decisions made.”* Furthermore Green et al. (2008, p. 13) describes the following characteristics of literature reviews:

- *“A clearly stated set of objectives with pre-defined eligibility criteria for studies;*
- *An explicit, reproducible methodology;*
- *A systematic search that attempts to identify all studies that would meet the eligibility criteria;*
- *An assessment of the validity of the findings of the included studies, for example through the assessment of risk of bias;*
- *A systematic presentation, and synthesis, of the characteristics and findings of the included studies.”*

Petticrew & Roberts (2006) argue that it is important to first discuss whether the systematic literature review is considered “narrow” or “broad” as it defines how the steps of reviewing process are carried out. A “narrow” systematic literature review addresses a very specific hypothesis or research question with specifics for the type of intervention or population (Petticrew & Roberts, 2006). Therefore a “narrow” systematic literature review more heavily excludes studies which don’t address these interventions or populations or simply don’t have the right study design (Petticrew &

Roberts, 2006). A “broad” systematic literature review more often includes various study design and addresses a more general research question (Petticrew & Roberts, 2006).

Tranfield et al. (2003) provided a roadmap (see table 8) to follow for performing systematic literature reviews. When comparing this roadmap with the approaches of Pettigrew & Roberts (2006) and Booth et al. (2012) there are quite some similarities. The approaches are compared side by side in table 8.

<b>Tranfield et al. (2003)</b>	<b>Pettigrew &amp; Roberts (2006)</b>	<b>Booth et al. (2012)</b>
Identification for the need for a review	A need for a systematic review	Planning and writing a literature review
Preparation of a proposal for a review	Refining the research question and boundaries	Defining the scope
Development of a review protocol	Deciding on the review's inclusion/exclusion criteria	Searching for literature and defining a protocol
Identification of research Selection of studies	The literature search	
Study quality assessment	Assessing study quality	Assessing the evidence base
Data extraction and monitoring progress	Synthesizing the evidence	Synthesizing included studies
Data synthesis		
The report and recommendations	Exploring heterogeneity and publication bias	Analysing the findings
Getting evidence into practice	Disseminating the review	Writing up and presenting data

Table 8: Similarities between approaches towards systematic literature reviewing

Beside the authors which were earlier mentioned Kitchenham (2004) also searched for different kinds of procedures or roadmaps which are used to perform systematic literature reviews as can be seen in table 9.

<b>Systematic Reviews Group</b>	<b>Australian National Health and Medical Research Council</b>	<b>Cochrane Reviewers Handbook</b>	<b>CRD Guidance</b>
			Identification of the need for a review. Preparation of a proposal for a systematic review
Define the question & develop draft protocol Identify a few relevant studies and do a pilot study; specific inclusion/exclusion criteria, test forms and refine protocol		Developing a protocol	Development of a review protocol
	Question Formulation	Formulating the problem	
Identify appropriate databases/sources. Run searches on all relevant data bases and sources. Save all citations (titles/abstracts) in a reference manager. Document search strategy.	Finding Studies	Locating and selecting studies for reviews	Identification of research Selection of studies

Researchers (at least 2) screen titles & abstracts. Researchers meet & resolve differences. Get full texts of all articles. Researchers do second screen. Articles remaining after second screen is the final set for inclusion			
Researchers extract data including quality data	Appraisal and selection of studies	Assessment of study quality	Study quality assessment
Researchers meet to resolve disagreements on data Compute inter-rater reliability. Enter data into database management software		Collecting data	Data extraction & monitoring progress
Import data and analyse using meta-analysis software. Pool data if appropriate. Look for heterogeneity	Summary and synthesis of relevant studies	Analysing & presenting results	Data synthesis
Interpret & present data. Discuss generalizability of conclusions and limitations of the review. Make recommendations for practice or policy, & research	Determining the applicability of results. Reviewing and appraising the economics literature	Interpreting the results	The report and recommendations. Getting evidence into practice

**Table 9: Similarities between approaches towards systematic literature reviewing as found by Kitchenham (2004)**

As we can see in the seven different approaches from table 8 and 9 there are many similarities between the steps that they take to perform a systematic literature review. When condensing the steps from all these approaches down to one bottom line the following can be said on the necessary steps that need to be taken.

1. The need for a systematic literature review needs to be clear
2. Defining a scope and research question accompanied by boundaries
3. Developing a review protocol with in-/exclusion criteria
4. Identifying and selecting research
5. Assessing the quality of the studies found
6. Extracting and synthesizing the data
7. Analysing, reporting and discussing the findings.
8. Presenting the findings

### **Strengths**

- Systematic literature reviews seek to draw all available literature on a topic together including both quantitative, qualitative or mixed studies (Grant & Booth, 2009) (Wickremasinghe et al., 2015)
- Makes use of clear and documented inclusion and exclusion criteria and is therefore reproducible (Wickremasinghe et al., 2015)
- Makes use of 'critical appraisal' which systematically considers factors as validity, generalizability and the used methods of included literature (Booth et al. 2012).

### **Weaknesses**

- By critically appraising studies and usage of inclusion and exclusion criteria, systematic literature review might offer a narrow view or only address a narrow research question, thus systematic literature reviews might not be suitable to more complex subjects or situation (Grant & Booth, 2009) (Malidou, 2014) (Wickremasinghe et al., 2015).

### 2.3.2 Meta-analysis literature reviews

According to Grant & Booth (2009) a meta-analysis literature review combines the quantitative results of studies included to provide a more precise image of the measured effect within these studies. Malidou (2014) argues that a meta-analysis review has a clear, transparent and replicable statistical method for analysing the included literature. For this type of literature review to be valid however, it needs the included studies to be quite similar in terms of: characteristics of the population being studied, the intervention that is being used to measure an effect, etc. (Grant & Booth, 2009).

<b>SALSA elements</b>	<b>Grant &amp; Booth (2009)</b>	<b>Malidou (2014)</b>
Search	Aims for exhaustive comprehensive searching for all literature	A systematic search strategy
Appraisal	Quality assessment which uses inclusion or exclusion criteria	Formal quality assessment with specified inclusion/exclusion criteria
Synthesis	Graphical or tabular with narrative commentary	Graphical or tabular with narrative commentary
Analysis	Numerical analysis of an effect across the studies found	Numerical (quantitative) analysis of an effect

Table 10: SALSA elements of a meta-analysis literature review

#### Strengths

- The strength of a meta-analysis literature review lies in its ability to combine statistical evidence from multiple studies into a more significant and complete image about a certain topic (Grant & Booth, 2009).
- Meta-analysis literature reviews also include small or inconclusive studies with statistical data (which might be excluded by other types of literature reviews) as these studies can still contribute towards complete image of measuring an effect (Grant & Booth, 2009).
- Meta-analysis literature reviews can settle existing controversies on a certain topic when two (or more) empirical studies have conflicting results by taking all statistical evidence into account (Malidou, 2014).

#### Weaknesses

- The weakness of a meta-analysis review lies in its need for the included studies to be similar as discussed by Grant & Booth (2009). Critics of meta-analysis literature review argue that sometimes studies may not be similar enough, thus combining 'apples and oranges' which may lead to a biased interpretation of the measured effect (Grant & Booth, 2009) (Malidou, 2014). Although this criticism can also be seen as a poorly performed meta-analysis literature

review, because this type of literature review requires all the decisions and criteria being made and used to be fully transparent (Borenstein, Hedges, Higgins & Rothstein, 2009)<sup>5</sup>.

### 2.3.3 Rapid Literature reviews

According to Grant & Booth (2009) this type of literature review was first seen as an unwelcome concession towards a systematic literature review, as time constraints hinders a systematic approach. But with the introduction of 'Rapid Evidence Assessment'<sup>6</sup> which offers a formal form of appraisal it gained legitimacy among researchers (Grant & Booth, 2009). Watt, Maddern, Cameron, Sturm, Babidge, Facey, Hailey & Nordehaug (2008)<sup>7</sup> defines a rapid literature review to be any systematic literature review which took between one and six months to produce with constraints on comprehensive search for literature.

SALSA elements	Grant & Booth (2009)	Malidou (2014)	Wickremasinghe et al. (2015)
Search	Completeness of the search is determined by time constraints	Complete search but with time constraints	Time-constrained search
Appraisal	Formal quality assessment limited by time	Time-limited quality assessment of the literature	Rapid evidence assessment
Synthesis	Typically narrative and tabular	Narrative and tabular	Narrative and tables
Analysis	Assesses quantities and quality of literature and identifies the direction of an effect within the selected literature	Maps quantity and quality of the literature and identifies the direction of an effect within the selected literature	Aimed at finding key-issues on a certain topic which might provide new research questions for more in-depth literature reviews

Table 11: SALSA elements of a rapid literature review

#### Strengths

- A rapid literature review provides a way for researchers to perform a quick literature review which is also replicable by following the reviewing process of a systematic literature review (Grant & Booth, 2009) (Malidou, 2014).
- Decisions on concessions being made are described and accounted for (Grant & Booth, 2009).

#### Weaknesses

- Shortening of the search process might lead to publication bias by omitting certain sources because of time constraints (Grant & Booth, 2009) (Malidou, 2014) (Wickremasinghe et al., 2015).
- Shortening of the appraisal of studies makes rapid literature review more prone to bias (Grant & Booth, 2009) (Wickremasinghe et al., 2015).

<sup>5</sup> Borenstein et al. (2009) extracted from Malidou (2014)

<sup>6</sup> An example of Rapid Evidence Assessment is given on this site

[http://media.wix.com/ugd/dded87\\_25658615020e427da194a325e7773d42.pdf](http://media.wix.com/ugd/dded87_25658615020e427da194a325e7773d42.pdf) (retrieved 01-06-2017)

<sup>7</sup> Watt et al. (2008) extracted from Malidou (2014)

- Shortening the synthesis process can result in overlooking contradictions or inconsistencies between included studies (Grant & Booth, 2009).

### 2.3.4 Umbrella literature review (review of reviews)

According to Grant & Booth (2009) the need for umbrella literature reviews occurred when systematic literature reviews became more common. Essentially a umbrella literature review is compiling evidence and results from systematic literature reviews into one overview (Grant & Booth, 2009)(Malidou, 2014)(Paré et al. 2015). According to Paré et al. (2015) umbrella literature reviews compare literature reviews with similar research questions in order to find contradictions between there results which may start discussions among researchers, policy makers and practitioners.

<b>SALSA elements</b>	<b>Grant &amp; Booth (2009)</b>	<b>Malidou (2014)</b>	<b>Wickremasinghe et al. (2015)</b>
Search	Identification of literature reviews only (no primary studies)	Systematically searching for systematic literature reviews	Includes existing literature review, preferably systematic studies
Appraisal	Formal quality assessment within the studies or the umbrella review itself	Formal quality assessment using an instrument for systematic literature reviews	Quality appraisal of each individual review is needed since quality of the reviews can differ
Synthesis	Graphical and tabular with narrative commentary	Narrative synthesis but may include graphical or tabular elements	Narrative, graphics and tables
Analysis	Aimed at creating a total image of knowledge on a broad topic, providing recommendations for practice and research	An overview of reviews which provides recommendations for practice, policies and future research	Provides an overview of reviews of which a conclusion or statement can be drawn on a broad topic

Table 12: SALSA elements of a umbrella literature review

#### Strengths

- Umbrella literature reviews offer a solution for researchers who are deciding on whether to perform one very broad literature review on a research area (at the cost of detail of individual findings)or a succession of heavily focussed reviews (which creates a fragmented total image of the research area) (Grant & Booth, 2009).

#### Weaknesses

- In many research areas there are not enough systematic literature reviews in order to carry out a umbrella literature review (Grant & Booth, 2009)(Malidou, 2014).
- Specific guidelines on how to conduct a umbrella literature review are not available making it hard for researchers who are not familiar with umbrella literature reviews to conduct one (Malidou, 2014).

### 2.3.5 Overlapping or other types of systematic literature reviews

#### Evidence paper

According to Wickremasinghe et al. (2015) an evidence paper is an overview of accessible peer-reviewed and grey literature. It differs from an overview (see sub-section 2.2.7) in that it does pose a critical quality appraisal of the evidence (Wickremasinghe et al., 2015). Synthesis is conducted in a

narrative and/or tabular way and the analysis is aimed at providing a potential base for a full systematic literature review or as a basis of evidence for policy makers (Wickremasinghe et al., 2015).

### **Realist literature review/Qualitative systematic literature review**

According to Paré et al. (2015) realist literature reviews were developed to enhance, extend or supplement systematic literature reviews. A main criticism on systematic literature reviews is that are conducted under simplistic or positivist assumptions like: if X is applied, Y will occur (Paré et al., 2015). This approach might work for research areas like medicine or education, but not for more complex research areas with more uncontrolled variable like social sciences (Paré et al., 2015). Realist literature reviews are aimed at explaining these more complex phenomena (Grant & Booth, 2009). The approach for synthesizing the found literature is narrative, sometimes accompanied by tables (Malidou, 2014). Analysis is focussed on finding and describing contradictions between studies which might be overlooked when performing a systematic literature review (Grant & Booth, 2009)

### **Systematic search and review**

According to Grant & Booth (2009) a systematic search and review combines a critical literature review (see sub-section 2.2.1) with a well-defined and comprehensive search strategy. Typically this type of literature review addresses a quite broad research question, enhancing the scope on topic in a research area which might give a more complete image than systematic literature review on the available literature (Grant & Booth, 2009). However, systematic search and review do not conduct any form of quality assessment (inclusion or exclusion criteria) making it prone to bias for supporting the researchers hypothesis/opinion (Grant & Booth, 2009).

### **Systematized literature review**

According to Grant & Booth (2009) a systematized literature review attempts to include elements of a systematic literature review and are conducted by those who are not able to make use of the resources necessary to carry out a full systematic literature review.

## **2.4 Differences between systematic and non-systematic literature reviews with regard to bias**

This section discusses what biases can be encountered when performing literature reviews. First a taxonomy and definitions of different types of biases are given. Secondly implications for the different kinds of literature reviews (see section 2.3) are described which will offer a base to describe the limitations of the non-systematic case study, as well as the repeat of this case study in a systematic fashion.

### **2.4.1 A taxonomy of bias**

According to Petticrew & Roberts (2006) bias refers to a wrong estimation of an effect which is either an over- or under-estimation. Booth et al. (2012) refers to bias as a systematic error within studies or reviews which may lead to wrongful conclusions about interventions, programmes or policies. Referring back to chapter one bias was discussed in two fashions; bias when selecting studies for a literature review and bias when interpreting the result for the selected studies. Song, Parekh, Hooper, Loke, Ryder, Sutton, Hing, Kwok, Pang & Harvey (2010) as well as Booth et al. (2012) presented a taxonomy for different kinds of bias affecting literature reviews. The following table (13) shows the different types of bias which were proposed by authors.

<b>Sort bias</b>	<b>Definition</b>
Dissemination bias	This refers to whether a study is accessible to a reviewer in terms of the reviewer being able to identify the results of the study or not. If the reviewer is not able to do this the study might be omitted from the review even though it might pose significant information (Song et al., 2010)
Publication bias	Refers to the bias that occurs when only studies with significant results in a preferred direction are published and thus possible to include in the review. Studies with less significant results might not be published at all and therefore be omitted, leading to over-exaggeration of the measured effect (Booth et al., 2012) (Egger & Davey-Smith, 1998) (Song et al., 2010).
Outcome reporting bias	Refers to publication bias within a study, when multiple outcomes are measured only the significant ones are reported in the study (Booth et al., 2012) (Song et al., 2010)..
Funding bias	Occurs when studies on, for example products, produce favourable outcomes for the company that makes the product because the company funds the study and its researchers (Booth et al., 2012) (Lexchin et al., 2003).
Time lag bias	Refers to studies with significant results are being published earlier than those with less significant results because in the studies with less significant results researchers might try to eventually find more significant results (Song et al., 2010).
Grey literature bias	Occurs when the results found within peer-reviewed studies are significantly different than those in conference papers, working articles, books, reports and dissertations (Booth et al., 2012) (Song et al., 2010).
Full publication bias	Refers to studies that are only being fully published when results are significant and in the right direction, otherwise only a partial study might be released leaving out crucial further research (Song et al., 2010).
Language bias	Occurs when a study in a foreign language (other than English) does not pose significant results (in the right direction) and is therefore not being translated and published in English. Reviewers who do not speak the language of the article might therefore omit it (Booth et al., 2012) (Song et al., 2010).
Multiple publication bias	Studies with significant results are more likely to generate multiple publications which thus can lead to an over-representation of one study within a literature review (Booth et al., 2012) (Song et al., 2010).
Place of publication bias	Refers to studies with more significant results and in a preferable direction are more likely to be published in more popular journals because of editorial choices of the journal or the readers' preferences (Song et al., 2010).
Citation bias	Occurs when a study is being cited because the direction or significance of the results supports the hypothesis of the researcher. This makes scanning through the reference list of a study less reliable as might overstate supportive studies, leading to the result of a literature review being skewed if those are included (Booth et al., 2012) (Song et al., 2010).
Database bias	Refers to the indexing choices being made by literature databases. For example, journals of third world countries might be under-represented as well as studies which are not in the language is favoured by the editors of the database (Booth et al., 2010)(Egger & Davey-Smith, 1998) (Song et al., 2010).
Media attention bias	Studies with strong results are more likely to be covered in newspapers or by radio/television/internet. This might lead to less significant studies being understated and represented (Song et al., 2010).
Interpretation	Refers to researchers (or reviewers) ability to synthesize, judge and weigh the

bias	results found in a study. Two researchers of different backgrounds might look at the same result in a different way thus drawing different conclusions based on their own background. This happens especially when the results are debatable or qualitative and therefore might lead to overstating certain results while other results are underrepresented (MacCoun, 1998).
------	---

Table 13: A taxonomy of bias

When looking at the given definitions of the different types of biases, certain terms are often mentioned by authors. Whether results of a study are significant or not and the direction of these results (positive or negative) seem to be a consistent factor among almost all types of bias. Song et al. (2010) acknowledges this relationship by stating that all types of biases mentioned in table 13, except for **funding bias** and **interpretation bias**, are linked to/associated with **publication bias**. Song et al. (2010) therefore uses a umbrella term for describing this: **publication and associated biases**.

#### 2.4.2 Differences between systematic and non-systematic literature reviews

In section 2.3 descriptions were given of the different types of both systematic and non-systematic literature reviews. Along with those descriptions came strengths and weaknesses for all different kinds of literature reviews. These descriptions and strength/weakness analyses implied bases for potential bias which could affect the outcome of the literature review. As a literature review is a study in which evidence of other studies is synthesized the found biases in sub-section 2.4.1 also hold true for literature reviews (Booth et al., 2012). However a specific link between certain types of literature reviews and specific types of biases are not strong or even non-existent.

When looking at the differences between systematic and non-systematic literature reviews (see section 2.3) it becomes clear that all systematic types of literature reviews require that the methods which are being used are clearly stated and documented which makes systematic literature reviews easier to reproduce. Another big difference between systematic and non-systematic literature review is the quality appraisal of studies found. More often do systematic literature reviews include some kind of formal quality assessment of the studies found whereas non-systematic literature reviews limit the use or do not have to include such practices (see tables 2-7). While according to Massaro et al. (2016) systematic literature reviews might be subjected to bias such as mentioned in table 13, they are transparent in their methods making potential sources of bias easier to identify and possibly deal with when reproducing the review. Non-systematic literature reviews do not have the obligation of being transparent about the methods used for selecting studies making the potential sources of bias possibly untraceable.

The absence of rigour in being transparent about the methods used for selecting studies to include into the literature review makes non-systematic literature review prone to any kind of bias associated with selection as readers cannot be certain that the review dealt with issues like publication and associated biases. Although the reviewer might have dealt with these issues it cannot be assumed and there has to be investigated on a case-by-case selection. Systematic literature reviews are also prone to the same types of biases, but these are detectable when looking at the methods for selection which are being used. Therefore, the reader can identify the shortcomings (biases) of the systematic literature review and can reviewers possible improve their systematic literature reviews by dealing with found sources of bias. An example can be given when looking at rapid literature reviews which are essentially short-cut systematic literature reviews (Grant & Booth, 2008).

*“Because of the constraints in resources the reviewers may have restrict to only using “Web of Science” as a database for their literature review which might be biased towards indexing studies that are written in French, while the review is about local French cheeses. This might pose database bias as it is logical that French researchers more often have researched French cheese than researchers from other countries. But due to “Web of Science” not favouring French studies a lot of work from those local French researchers might be omitted which might provide a skewed literature review”.*

By clearly stating that due to resource constraint only “Web of Science” was used, the reader or future researcher can deal with or take into account the potential issue of database bias. The implication which can be given is that the found types of publication and associated biases are applicable to both systematic and non-systematic literature reviews dependent on the choices that are made by the reviewer(s).

According to MacCoun (1998) interpretation bias is a commonly found phenomenon in literature with many motivational, intentional and purely cognitive reasons. However MacCoun (1998) describes the effect of interpretation bias as being relatively small and subtle. Even when the intentions of, for example a reviewer, are motivated by finding evidence to support a hypothesis they can be caught by using systematic empirical research methods (MacCoun, 1998). MacCoun (1998) further argues that there will always be some room for interpreting results in a different way among researchers. The implication for literature reviews here is that the effect of interpretation bias is relatively small.

### 3 Methodology

This chapter discusses the way in which the research question and sub-questions are going to be answered in detail. This chapter first discusses the selection of a case study (non-systematic literature review) on the Balanced Scorecard which is used to compare differences between the results of non-systematic and systematic literature reviewing methods. Secondly, the method of recreating the case study in a systematic manner is discussed. At last the manner in which the repeat review is compared to the case study is explained.

#### Chapter Summary

- After a systematic search in both Scopus and Web of Science for a non-systematic literature review Madsen (2015) was chosen as case study.
- The repeat review which is conducted is a rapid literature review, which uses the methods and steps of a systematic literature review even though compromising some elements to meet the available resources (time, database access, etc.).
  - The aim, scope, research question and boundaries of the repeat review are based on the case study of Madsen (2015).
  - A review protocol accompanied by in- and exclusion criteria is provided by using the PICOC criteria in order to select studies.
  - Boolean Logic is applied to search more effectively for relevant studies to include in the repeat review.
  - Formal quality appraisal is not applied on the repeat review, but the included studies' journals are examined.
  - The study designs of the included studies are examined as well as the relatedness between the included studies by means of a network analysis.
  - A thematic analysis is applied based on the coding of information found in the included studies.
- Comparison between the repeat review and the case study consist of comparing the examination of journals, comparing the study designs of the included studies of both reviews as well as the network analyses and finally thematic analyses are compared. Out of these comparisons differences are derived and sources of bias.

#### 3.1 Selecting and analysing a case study

This section discusses the subject of the case study (Balanced Scorecard) and how the case study is selected.

##### 3.1.1 Background on the subject of the case study

As earlier mentioned in section 1.4 the subject of the case study is being given by this projects supervisor which is the Balanced Scorecard. This subject was chosen, because according to Bassioni et al. (2004), Marr & Schiuma (2003) and Neely (2005) the research area of performance measurement systems is highly dominated by the Balanced Scorecard. Thus meaning that it is an important subject within the research area. Neely (2005) furthermore argues that researchers in this area are looking to the Balanced Scorecard according to proven practices at organizations and are looking for theoretical explanations as to how and why these practices of the Balanced Scorecard

work within organizations. This offered ground for researchers to perform literature reviews on the Balanced Scorecard which makes it a suitable subject for the case study.

Kaplan & Norton (1992) first introduced the term: “Balanced Scorecard” or “BSC” and thus defines the starting point of the topic within the research area of performance measurement. The BSC is a tool with which companies can translate their strategic objectives into performance measures. Kaplan & Norton (1992) developed four different aspects for which they formulated goals and measurements which can be seen in figure 2 and are explained based on Kaplan & Norton (1992).

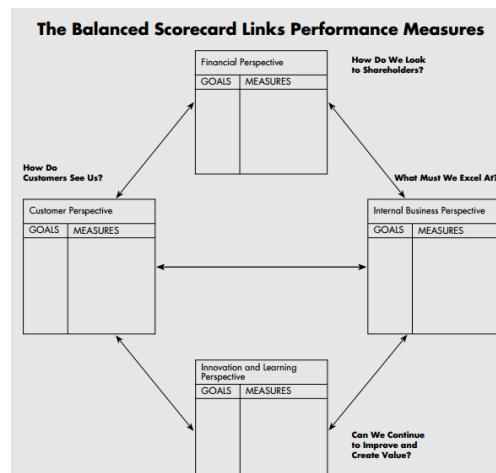


Figure 2: The Balanced Scorecard (BSC) extracted from Kaplan & Norton (1992)

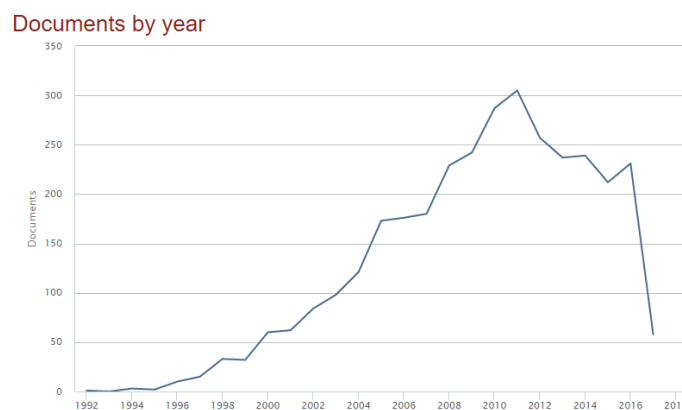
The first is the *financial perspective* in which an organization could ask itself the following question: “How do we look to shareholders?” (Kaplan & Norton, 1992). Examples of goals given were to survive, to succeed and to prosper financially, which could be measured by the cash flow, market share growth and quarterly increase of sales (Kaplan & Norton, 1992). The second perspective is the *customer perspective* in which an organization could ask itself the following question: “How do customers see us?” (Kaplan & Norton, 1992). Goals in this perspective are the amount of new products made and customer partnership which can be measured by the on-time delivery count of products and the percentage of sales from new products (Kaplan & Norton, 1992). The third perspective is the *internal business perspective* in which an organization could ask itself the following question: “What must we excel at?” (Kaplan & Norton, 1992). Goals for this perspective are raising the technological capability and manufacturing excellence which can be measured by the cycle time and unit cost (Kaplan & Norton, 1992). The fourth and last perspective is the *innovation and learning perspective* in which an organization could ask itself the following question: “Can we continue to improve and create value?” (Kaplan & Norton, 1992). Goals here are product focus and time-to-market which can be measured by the time needed to introduce the next generation of products and the percentage of products which equals 80% of the sales. (Kaplan & Norton, 1992)

### 3.1.2 Finding a case study

Selecting the right non-systematic literature review for comparison will be crucial for the rest of the research project as it will set some boundaries for the recreated systematic literature review that will be performed later on. One of the first steps in this process is to collect all the existing non-systematic literature reviews on the balanced scorecard. For doing this both article search engines:

Scopus and Web of Science are used (Sources as EBSCOhost and others are omitted because they are not available for the researcher to be used).

The first step in finding the literature reviews on the BSC was to enter the string: “Balanced Scorecard” into the search field (1-3-2017) for finding it in articles’ titles, abstracts and keywords in Scopus. Next the same is done for Web of Science where it is searched as a topic string. Scopus presented 3408 found document on the matter of the BSC while Web of Science found 2602. The next step was to filter out the reviews from these piles of document which was done by checking the review filter in both search engines. This led to Scopus finding 248 reviews concerning the BSC while Web of Science found 57 reviews on the matter. The following step was to exclude the review which are done in another language than English as only the English ones can be analysed due to the restriction in the researchers’ multi-lingual abilities. This led to 216 reviews in Scopus and 48 in Web of Science. The next important step is to look at the time in which the literature review was made. The first article to ever appear on the BSC was that of Kaplan and Norton in 1992 which means the research area on this topic is quite young. Since that article many authors have followed into this research area which resulted in the following graph on how many articles there are written on this subject every year.



**Figure 3: Number of articles concerning the Balanced Scorecard (Source: Scopus, retrieved 1-3-2017)**

As we can see in figure 3 there has been an increase in articles since the early 2000's. If a literature review of the BSC would be used which was made in the year 2000 it would also mean that the majority of “today's” knowledge base wouldn't be included. This will result in a perhaps well performed literature review being made, but unfortunately outdated. Therefore the next step for refining our results is to exclude review which have been made before 2015. The reason for this is to ensure that at least 80% of the knowledge base of the BSC is possibly being included into the case study. This leads to a set of 13 literature reviews in Scopus and 13 literature reviews in Web of Science. Since we are looking for a non-systematic literature review on the balanced scorecard, the systematic ones are being omitted by means of an analysis of the abstracts. Furthermore, studies that do address the BSC but do not use it as the primary subject of their review are also omitted. This lead to a final cut of 5 non-systematic literature reviews in Scopus and 4 in Web of Science. Because the case study is being fully analysed it is necessary to have access to the full-text of the study. Since the University of Twente does not have access to every journal it is necessary to restrict only to the studies in journals which the University of Twente does have access to. This leads to only one study which is a non-systematic literature review that is fully accessible and was published in 2015. The

study of Madsen (2015) will therefore serve as the case study and will be thoroughly analysed in the following section.

### 3.2 Conducting the repeat review

This section describes the manner in which the selected case study is going to be analysed further by means of repeating it in a systematic manner. Therefore the rigour of the systematic literature review as described in sub-section 2.3.1 is used. However considering the boundaries of this research project (short time frame and one unexperienced reviewer) it is necessary to make some concessions towards this rigour thus becoming more of a **rapid literature review** (see sub-section 2.3.3). The steps of a systematic literature review are all described with their concessions, if any.

#### **The need for a systematic literature review**

As this is a repeat of the case study of Madsen (2015) already provided a need for a literature review in the problem statement. This need is therefore also used for the systematic repeat of the case study.

#### **Defining a scope and research question accompanied by boundaries**

The scope of the repeat review is also derived from the study of Madsen (2015). Since Madsen (2015) does not provide a research question but only an aim a research question is developed based on the provided aim. For the purposes of the repeat review it is also necessary to provide boundaries to this research question which are provided along.

#### **Developing a review protocol with in-/exclusion criteria**

When developing in-/exclusion criteria Petticrew & Roberts (2006) argue that it is important to first discuss whether the systematic literature review is considered “narrow” or “broad” (see sub-section 2.3.1). Defining whether the systematic literature review is “narrow” or “broad” will be based on the aim, research question and boundaries which were defined earlier. This in turn provided a base for deciding on the in-/exclusion criteria for studies found. These criteria are implemented in a review protocol which is important to exclude studies which are not concerned with the subject. Petticrew & Roberts (2006) and Booth et al. (2012) provide a tool (PICOC) for describing such a protocol with in-/exclusion criteria.

- **Population:** What or whom is going to be studied?
- **Intervention:** What or which effect is being studied which affects the population?
- **Comparison (optional):** What is the alternative for the performed intervention? (such as a non-intervention)
- **Outcomes:** What are the outcomes which are going to be measured?
- **Context:** What are the settings and boundaries in which the PICO criteria are measured? (timeframe, etc.)

Every element of PICOC is described providing the needed in-/exclusion criteria for finding the right studies to include in the repeat review

#### **Identifying and selecting research**

In performing a systematic literature review it is important to systematize the searching for relevant studies as well. According to Booth et al. (2012) it is first important to derive search terms or concepts from the research questions. As earlier mentioned the main themes in the research

question are focused on the usage of the BSC in SMEs. BSC and SMEs are therefore the core concepts on which the search for studies are focussed. Adjacent term for the BSC (such as performance measurement) and SMEs (such as medium sized firms) are defined in order to find studies which use other terminology than Madsen (2015)

Boolean logic is used to appropriately combine different concepts and terms to search. Boolean logic is a way of adding, subtracting and multiplying search terms to reach a more complete total of search to find what you are looking for (Booth et al., 2012). Boolean logic uses the operators **AND**, **OR** and **NOT** (Booth et al., 2012).

- **AND** is used to combine different concepts together, thus narrowing the search.
- **OR** is used to combine different terms within the same concept, thus expanding the search.
- **NOT** is used to exclude irrelevant search terms, thus narrowing the search.

Using these operators to combine the concepts will provide a more complete search within the databases of Scopus and Web of Science. As earlier mentioned these two databases will be the only source of the studies included for this repeat review as time is limited.

After the search strings are all used it is important to take a closer look at the studies which have been found. Given that the search strings should provide mostly relevant studies it is possible that they are not relevant but do fit the search strings. For example a study might discuss the BSC and does mention SMEs but does not discuss the usage of the BSC in SMEs but rather mentions SMEs together with large firms in a more general sense. Studies might also be omitted due to the language in which they are written, which poses a problem to the reviewer. These sifts are described accompanied by the number of studies included or excluded after every sift. No secondary searching is performed by means of searching through references of the primary found studies out of Scopus and Web of Science, because of time constraint of the reviewer.

### **Assessing the quality of the studies found**

After finding the eligible studies for inclusion in the repeat review, their quality should be assessed. Booth et al. (2012) and Petticrew & Roberts (2006) both agree that there may be over a hundred different ways to appraise the quality of studies, especially for qualitative studies. Booth et al. (2012) suggests that the use of checklists is appropriate for novices to use for assessing the quality of studies. One of those checklist is the CASP (Critical Appraisal Skills Programme) list<sup>8</sup> which uses a set of ten questions to appraise a study's quality. For example this list contains the question: "Was the research design appropriate for to address the aim of the study?" Which can be answered with yes, no or cannot tell. The more questions which are answered with yes the higher the quality of study is appraised and therefore the more trustworthy the results of that study are. However, Dixon-Woods Booth & Sutton (2007) found during their study that six different reviewers came up with six different quality appraisals of the same study. This shows that the checklists are prone to the subjective interpretation of the reviewer. Given the lack of experience of the reviewer performing this repeat review the results of the quality appraisal might not be correct as well. Booth et al. (2012) also states that performing quality assessment is a very time consuming process which preferably has to be performed by multiple reviewers, both of these resources are not available. Therefore this

---

<sup>8</sup> Source: [http://media.wix.com/ugd/dded87\\_25658615020e427da194a325e7773d42.pdf](http://media.wix.com/ugd/dded87_25658615020e427da194a325e7773d42.pdf) retrieved 01-06-2017)

repeat review will lack in a systematic quality appraisal, however when synthesizing the results from the studies, some of the checklist's question are used to put the results into perspective.

What the resources do allow the reviewer to do is assess the impact factor of the journals of which the selected studies came from. The impact factor of a journal is determined by dividing the number of citation towards the journal by the total number of publications of the journal between two years (Institute of Scientific Information, 1994). The higher the number of citations of a journal is compared to its publications the more impact the journal has. This assessment is made in order to find a potential source of place of publication bias (see sub-section 2.4.1).

### **Extracting and synthesizing the data**

The "extraction" of data refers to which information from the included studies in the repeat review is being used and how. Tranfield et al. (2003) and Booth et al. (2012) both describe the use of forms on which the relevant information of a study is noted systematically, making it easier to process and analyse later on. The extraction form (see appendix A) consists of a list including the authors, year of publication, journal, key words, type of study and the abstract. This extracted data is used during the synthesis and analysis of the results. The data synthesis is first presented in tabular forms, wherein the descriptive statistics are shown about the types of studies included in the reviewing process. Also the origin of the studies are shown in tables to determine whether most knowledge comes from one or a few journals or not. Also the results are briefly mentioned in tabular form, with the full results and relevance to the subject noted in the extraction forms.

The references of each study will also be exported into an .CSV Excel list. After extracting the references of the studies this .CSV file can be imported in a free analysis application called VOS viewer. This analysis application will offer an analysis of the data found which makes it more likely to find patterns and biases within the research area. VOS stands for visualization of similarities and is developed by Waltman, Van Eck & Noyons (2010). VOS viewer combines an optimization and a clustering algorithm for creating an unified approach for mapping and clustering bibliometric networks (Kovacs, Van Looy & Cassiman, 2014). VOS viewer can form an image in which clusters can be shown for various types of research. In the case of this research project it means that clusters are formed of the studies that have been extracted. VOS optimization algorithm ensure that the distances between articles and clusters on the visual image represent the actual relatedness between these articles and clusters (Kovacs et al., 2014).

For creating the least biased repeat review on the usage of the BSC in SMEs an examination of the foundation of this research area is performed. For analysing the foundation of the BSC research area the bibliographic coupling technique is used. This will show which authors dictated the research field in the past which possibly indicates a bias when just a few authors form the foundation. According to Kovacs et al. (2014) this technique will couple two articles when they share one or more references to the same studies. When many studies refer to the same studies in their references, clusters will form showing a common foundation. When a research area is heavily dominated by a certain author or authors, there might be **citation bias** (as mentioned in sub-section 2.4.1). Big clusters point to a large common foundation on which included studies of the repeat review and the case study are built, which indicates citation bias. To summarize, first a tabular synthesis is performed on the data from the extraction forms, secondly a graphical synthesis is performed by means of a clustering map from VOS viewer on the references from the included studies.

### Analysing, reporting and discussing the findings

After the data is synthesized it needs to be analysed in order to draw up results. Madsen (2015) used a thematic approach for analysing the data by mentioning beforehand what those themes were. The repeat review also uses a thematic analysis, but uses the results from the study as a starting point. Booth et al. (2012) mentioned **coding** of the results as a technique for finding themes among the studies. According to Booth et al. (2012) this first involves free-line-coding which refers to highlighting certain words or sentences within the results of the studies related to the subject. For example:

Study results	Code
SMEs don't understand how to use the BSC to support their strategy planning process.	Knowledge Benefits

Table 14: Free-line-coding (example)

With these codes of the results general themes are likely to surface as some codes are predominantly present among the results of different studies. This indicates that the code is a general, and thus important, theme among the studies. Sometimes these codes are present together with other codes meaning that they together represent a theme among the results. After finding these themes a narrative analysis is used to describe what the results of the included studies say about the themes found. At this stage it is important, according to Booth et al. (2012) to describe the directions of the results and discussing whether there is a general agreement among the studies with regard to the direction of the results. After confirming the direction of the results of themes, the themes are described which offers a base to identify research gaps and form a research agenda.

### Presenting the findings

This last step concerns the way in which the repeat review is performed toward the general or specific public. Since it is accurately described how the repeat review is performed and the results are presented in chapter 4, there is no need to further explain how the repeat review is structured.

## 3.3 Comparing the case study to the repeated study

Referring back to the research goal, this section discusses the manner in which the original non-systematic study of Madsen (2015) is compared to the repeated systematic study (Essentially comparing section 3.2 to section 3.3). More specifically this section shows how sub-research question 3 and 4 are answered.

In order to answer sub-research question 3 (*“What differences do occur when selecting studies for the literature review between a non-systematic literature review and a systematic literature review?”*) the search and selection strategy of both the case study and the repeat case study needs to be compared. Since Madsen (2015) does not specify how the literature was selected we only have the actual references of Madsen (2015) for comparison. First the tabular and graphical analyses on the characteristics of the included studies are compared (author(s), year of publication, journal, etc.) to spot differences. Secondly, the impact factor of the journals of the studies included by Madsen (2015) are assessed and compared to those of the repeat review in order to potentially find place of publication bias (see sub-section 2.4.1). After that the analysis with clustering images (from VOS viewer) becomes important. The references of Madsen (2015) are also mapped which shows how the foundation of Madsen's (2015) literature review is structured. This clustering map will also be compared with the clustering map of the repeat review to find potential differences. The found

differences provide a base for a discussion on how the reviewers choices (subjected to bias) affected the selection of studies for inclusion in the review.

For answering sub-research question 4 (*“What differences do occur when interpreting the results from the selected studies material a non-systematic literature review and a systematic literature review?”*) the analysis of both the case study and the repeat case study needs to be compared. Since both studies use a thematic analysis for analysing the data the differences can be compared one-on-one. The differences between the themes will be discussed in a narrative fashion including the directions of the results within the themes (in case similar themes are found among both the case study and the repeat case study which shows results in different directions). The research agendas are also compared to find similarities and differences. The found differences show what the effect of the reviewers choices (subjected to bias) are on the actual product of a literature review both systematic and non-systematic.

The comparison between the case study and the repeated case study are discussed in a separate section in chapter 4 as it important to show the differences and similarities between the non-systematic and systematic approach in literature reviewing side-by-side. This is done to ensure that the differences are as clear as possible in order to answer the main research question.

## 4 Results

This chapter starts by introducing the case study of Madsen (2015) and stating what Madsen (2015) aimed for in his study. This is followed by an analysis of the included studies by Madsen both tabular and graphical to find potential sources of bias. At last the thematic analysis of Madsen (2015) is summarized. Secondly the repeat review is presented following the methodology as presented in section 3.2. At last both reviews are compared to each other on their selection process and outcomes as well as their interpretation process and outcomes to find similarities and differences.

### Chapter Summary

- In the selection process of studies to be included for the repeat review, 6 potential studies out of 172 were excluded because they were not in English. This is a source of **language bias** affecting the potential studies for inclusion in the repeat review. (see sub-section 4.2.4)
- There were no significant differences to be found between both reviews when looking at the year of publication (see sub-section 4.3.1)
- Looking at the included peer reviewed studies for both the repeat review (30 out of 40) and the case study (28 out of 39) they both come from relatively high ranking journals posing a source of **place of publication bias**. Although only 8 out of 40 boasted this high average ranking in the repeat review and 7 out of 39 in the case study this type of bias still affects a roughly equally number of the included studies. (see sub-section 4.3.1)
- Looking at the included studies for both the repeat review and the case study they do have quite the same pallet of types of data collection. The reviewer was not able to identify the study design of one included study in the repeat review posing a slight sign of **dissemination bias** (see sub-section 4.2.6).
- When comparing the authors of the included studies for both reviews the case study (of Madsen (2015)) did refer to his own work 4 times. Moreover these 4 included studies were all based on the same data which poses a potential source of **multiple publication bias** (see sub-section 4.2.6).
- The network analysis on the included studies for both reviews revealed 6 clusters (consisting of 3 to 6 studies) with a total of 26 articles in the repeat review and 3 clusters (consisting of 3 to 10 studies) with a total of 19 articles in the case study. The impact of **Citation bias** in this case is hard to determine but is present in both the case study and the repeat review (see sub-section 4.3.1).
- The thematic analysis of the included studies of both review shared similar themes and many similar findings since both reviews share 11 of the included studies. However many differences between both reviews were also present, even among the interpretation of information in the shared included studies which poses a potential source of **interpretation bias**. (see sub-section 4.3.2).

### 4.1.1 Introducing the case study

To describe the aim of the case study of Madsen (2015) first a context is described in which it is situated. The title: "The Balanced Scorecard in the context of SMEs (Small/Medium Enterprises): A literature review." gives a first clue about the context of the study. Madsen (2015) begins by stating

that the BSC is one of the most influential management ideas of our time and that the BSC is widely adopted among large firms. Secondly Madsen (2015) describes that the BSC is has been increasingly subjected to academic inquiry and that there exists a large body of research on this subject. Madsen (2015) argues that this body of literature mainly focusses on the BSC in large firms and further states that the usage of the BSC in SMEs only has been studies by few scholars. Madsen (2015) further states some possible explanations as to why the usage of BSC in SMEs has been understudies such as: the BSC being designed for large firms and preferences of researchers to rather conduct research onto larger firm. This introduction leads Madsen (2015) into the following aim for his literature review:

*“Against the background outlined above, the current paper aims to add to the BSC literature in two ways. First, to take stock of the available literature about the use of the BSC in SMEs using a narrative review approach. Second, to identify underexplored areas and sketch an agenda for future research in the SME context. Although focusing on the BSC, the paper discusses issues which should also be of general relevance to research on the management accounting and control practices of SMEs.”*

When looking at the aim of Madsen (2015) it can be argued that its nature is exploratory as it sets out to find available literature on the BSC in SMEs and then use this to identify gaps in research to set up a research agenda for future researchers. Madsen (2015) even goes beyond this preliminary goal by stating that its literature review also offers implications in a broader sense, namely the area of management accounting and control practices within SMEs.

#### 4.1.2 Description of the selection process of the case study

Since Madsen (2015) does not provide a search and selection process only the included studies in the reviewing process are discussed. Madsen (2015) included a total of 39 studies in his reviewing process.

##### Authors of the included studies

When looking at the authors of the included studies there are only three authors which have provided multiple contributions towards both reviews as can be seen in table 15.

Review	Multiple publications	
Author(s)	Kaplan & Norton	Madsen
Studies	- Kaplan & Norton (1996) - Kaplan & Norton (2004)	- Madsen (2014a) - Madsen (2014b) -Madsen & Stenheim (2014a) - Madsen & Slåtten (2015)

Table 15: Authors with multiple contributions to the case study

Looking at Kaplan & Norton they worked together on two studies which are included in the case study. Kaplan & Norton (2004) is a book which does refer to Kaplan & Norton (1996), but is not based on the results of Kaplan & Norton (1996) and uses over a hundred different sources to gather data from. Looking at the case study Madsen does include 4 studies which he (co)wrote during 2014 and 2015. Moreover, these 4 studies all used the same source of data referring to interviews with adopters of the BSC in Scandinavia in a study performed by Madsen himself in 2011 (Madsen, 2014a)(Madsen, 2014b)(Madsen & Stenheim, 2014a)(Madsen & Slåtten, 2015). This is a source of **multi-publication bias** as the same data is used multiple times which over exaggerates its effect in

the case study of Madsen (2015). The impact of this **multi-publication bias** on the study of Madsen (2015) is significant as it affects over 10% (4 studies out of 39) of his evidence base.

### Year of publication of the included studies

When looking at the included studies in the case study, Madsen (2015) does include studies of which some of them are quite old, while others are as recent as the year Madsen (2015) was published (see figure 4).

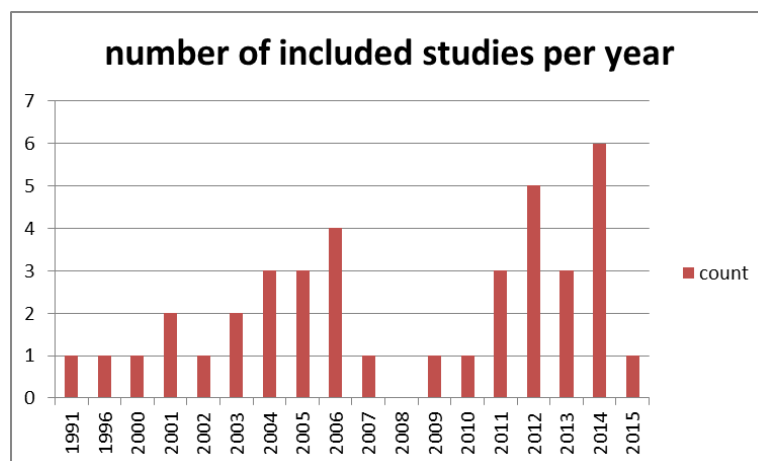


Figure 4: Year of publication of the included studies in the case study

Looking at figure 4 it becomes clear that there is an even distribution among studies which are over 5 years old (2009 and below) at the time of the review of Madsen (2015) and studies under five years old (2010 and up). According to Pautasso (2013) a literature review should always be up-to-date and thus look for more recent studies to include in the reviewing process. However older studies (sleeping beauties) should also be included to balance out a review, providing a more complete overview of the research area (Pautasso, 2013). Therefore Madsen (2015) according to logic of Pautasso (2013) used a well distributed set of included studies for his review.

### Type of data collection of the included studies

The studies included by Madsen (2015) show a great variety in which the data was gathered for reaching their results. The various types of data collection of these studies are classified in figure 5.

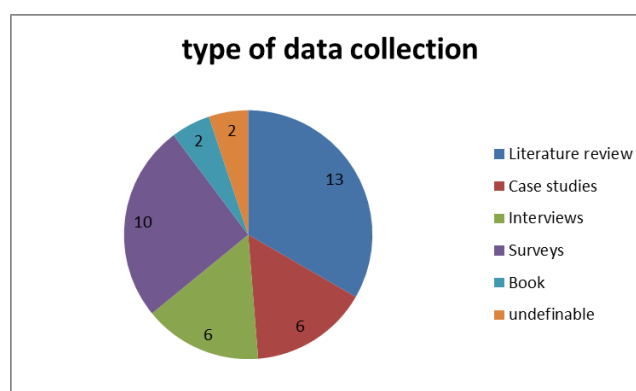


Figure 5: type of data collection of the included studies in the case study

When looking at figure 5 the “undefinable” stands for two included studies which were written in the Norwegian language, making it impossible for the reviewer to determine what type of data collection

was used. To assess whether the quality of these types of data collection is high or low Tranfield et al. (2003) uses of the hierarchy of evidence (see figure 6)

Hierarchy of evidence	
I-1	Systematic review and meta-analysis of two or more double blind randomized controlled trials.
I-2	One or more large double-blind randomized controlled trials.
II-1	One or more well-conducted cohort studies.
II-2	One or more well-conducted case-control studies.
II-3	A dramatic uncontrolled experiment.
III	Expert committee sitting in review; peer leader opinion.
IV	Personal experience.

Figure 6: Hierarchy of evidence (Tranfield et al., 2003)

From a positivistic stance most of the evidence of the included studies belongs to the second category within conducted cohort studies or case-control studies, because not one of the included studies uses randomized control trials of any kind. According to Evans (2002) case studies, interviews and surveys are all considered poor quality evidence. However Evans (2002) is relating this to the medical field where randomized control trials are more appropriate and easier to design than in the field of business research. Madsen (2015) therefore might not have access to higher quality evidence to use in his review and uses the best sources of evidence available to him.

### Impact of the included studies

Out of the 39 studies included by Madsen (2015) 29 were peer-reviewed and published in papers. Looking at the average cumulative impact factor (1,507, see appendix B) it can be considered quite high as compared to the average impact factor among journals within the subject area. Among the 1394 journals the average cumulative impact factor of both reviews rank on the 130<sup>th</sup> place<sup>9</sup>, which puts it the top 10%. This poses a source of **place of publication bias** as studies with less significant results are less likely to show up in popular journals (Song et al., 2010). However when looking at the impact factors of the individual journals it is clear that only 7 out of the 39 included studies in the case study have an impact factor above 1,500 (see appendix B). Journals such as Academy of Management Review have such a high impact factor (>8,000) that it skews the average while only one of the included studies in the review of Madsen (2015) came from this journal. This in turn decreases the impact of **place of publication bias** for the case study. Furthermore when looking to impact factors in general, 67% of all journals rank above 1,000 according to Journal Citation Report (Web of Science) lessening even more the effect **place of publication bias**.<sup>10</sup>

Looking at the individual journals which become apparent in the included study only the 'International Journal of Productivity and Performance Management' is noticeably present with 5 of the included studies being published by it (see appendix B). It seems unlikely that Madsen (2015) selected these studies for his review because they were published in this journal, since the impact score of 0,607 (see appendix B) is relatively low.

<sup>9</sup> Retrieved from [http://www.scimagojr.com/journalrank.php?area=1400&page=3&total\\_size=1394](http://www.scimagojr.com/journalrank.php?area=1400&page=3&total_size=1394) at 28-07-2017

<sup>10</sup> Retrieved from <http://mdanderson.libanswers.com/faq/26159> at 28-07-2017

## Network visualisation and case analysis

VOS viewer requires an output from either Web of Science or Scopus. In the case study Madsen refers to two studies (Braam & Heusinkveld, 2002)(Kjode, 2003) which were only presented at the university of which the authors were affiliated with. These studies cannot be found via Web of Science nor via Scopus and could therefore not be included in the network visualisation.

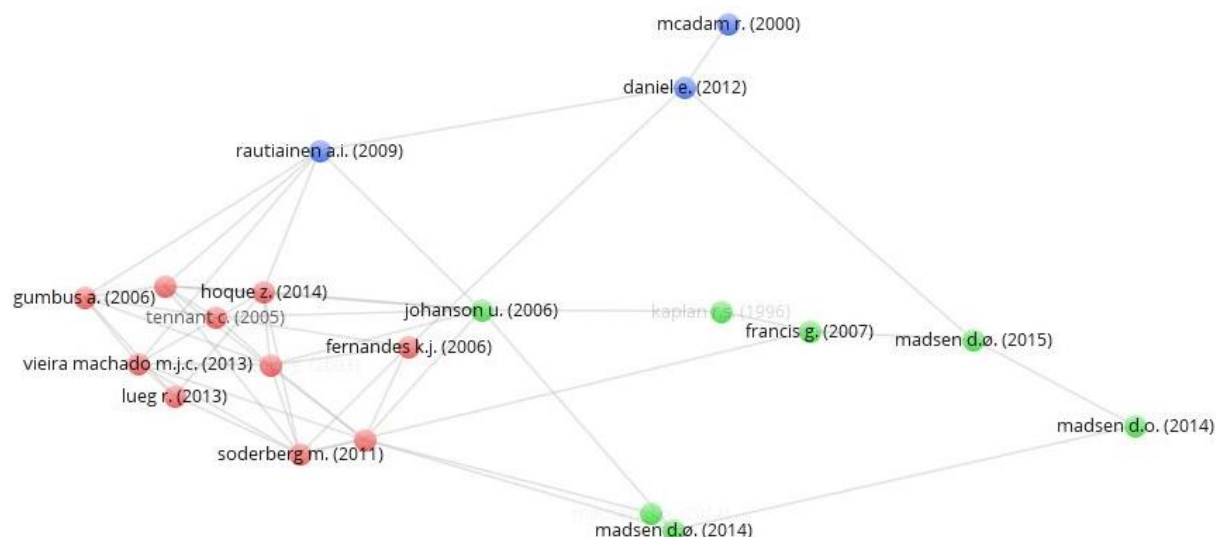


Figure 7: Network visualisation of the case study

As can be seen in figure 6 only three clusters were formed spanning a total of 20 out of the 37 articles (39 minus the two omitted studies). 17 of the included articles did not form a connection or cluster with one of the other included studies meaning that they don't share a common foundation. The studies within the clusters do share a common foundation and are described in table 16.

Cluster 1: implementation	Cluster 2: future research	Cluster 3: adoption
Fernandes et al. (2006)	Francis & Holloway (2007)	Daniel et al. (2012)
Gumbus & Lussier (2006)	Johanson et al. (2006)	McAdam (2000)
Hoque (2014)	Kaplan & Norton (1996)	Rautiainen (2009)
Lawrie & Cobbold (2004)	Madsen & Stenheim (2014)	
Lucianetti (2010)	Madsen & Slatten (2015)	
Lueg et al. (2013)	Madsen (2014a)	
Soderberg et al. (2011)	Madsen (2014b)	
Speckbacher et al. (2003)		
Tennant & Tanoren (2005)		
Machado (2013)		

Table 16: Cluster description of the case study<sup>11</sup>

Looking at the evidence from table 16 the case study of Madsen (2015) does include significant clusters of studies including 7,7% (3 out of 39 included studies) to 25,6% (10 out of 39 included studies) of the total amount. The studies found in cluster 1 are very closely related to each other as compared to the other two clusters which are both quite widespread. Cluster one therefore shows the biggest indicator of **citation bias** as these 10 studies are likely to share many references and thus a common foundation. Looking at the recurring themes in cluster 1 it becomes apparent that they all address the issue of implementation of the BSC in SMEs. The studies of cluster 2 are only used by

<sup>11</sup> The cluster names are based on an interpretation of the common theme by the reviewer after reading the included studies of a cluster.

Madsen (2015) for building his research agenda (except for Kaplan & Norton (1996)). Looking at cluster 2 it does become clear that there is a case of **citation bias** as Madsen (2015) cites his own work which is used to support his own research agenda in his literature review. The studies in cluster 3 all address the adoption of the BSC by SMEs.

#### 4.1.3 Interpretation of the information within included studies of the case study

##### **Thematic analysis**

Madsen (2015) starts by defining four themes which according to Madsen (2015) all concern SMEs specific BSC literature before starting his thematic analysis (the full thematic analysis can be found in the article of Madsen (2015)).

Knowledge and awareness: Madsen (2015) argues by means of included survey studies that SMEs have little knowledge about the BSC nor awareness. However Madsen (2015) does state a survey study which found that SMEs in Norway had come in contact with the BSC via conferences, and to a lesser extent, consultants.

Adoption and diffusion: Madsen (2015) argues that decision for adopting the BSC or not within SMEs is based on both internal and external factors. Madsen (2015) considers contact with consultants and visitations of conferences as an external factor influence the decision of SMEs. Madsen (2015) states that SMEs do investigate the usage of the BSC but in most cases reject it, because SMEs see the BSC as non-applicable. Madsen (2015) further argues that adoption rates of the BSC in SMEs are low by means of stating results from included survey studies.

Implementation: Madsen (2015) argues that implementation of the BSC in SMEs is generally faster than in large firms due to SMEs having a simpler organisational structure. SMEs further also tend to implement a simpler and earlier version of the BSC according to Madsen (2015). Furthermore Madsen (2015) states that when implementing the BSC in a SME the BSC should be adapted towards the circumstances and needs of the SME.

Experiences: Madsen (2015) starts by stating that the BSC can successfully be applied in SMEs, even avoiding pitfalls which larger organisations face when implementing the BSC. According to Madsen (2015) the main benefits of implementing the BSC in SMEs is that it makes management processes more effective in terms of setting strategic objectives while maintaining the competitive advantage of having a simple corporate structure. The main problems for using the BSC in SMEs are lack of resources (knowledge about the BSC, general lack of HRM department, etc.) and that the BSC is too rigorous to fit the relatively unstable environment of SMEs (Madsen, 2015). Moreover, Madsen (2015) specifically mentioned the lack of training and understanding of the BSC within SMEs to be a barrier.

##### **Research agenda**

After describing and reviewing literature among the four themes Madsen (2015) developed a research agenda about these subjects in a narrative manner.

- Knowledge and awareness: Madsen (2015) argues that future research should focus on why SMEs are not aware and have little knowledge about the BSC as it potentially explains the low adoption rates. Madsen (2015) also argues that future research should pay attention to how SME managers learn and come in contact with the BSC.

- Adoption and diffusion: Madsen (2015) argues that future research in this area should focus on why SMEs choose to adopt or not adopt the BSC in terms of various factors. The adoption rate in this case should also be further mapped among SMEs because there is not much evidence on this matter (Madsen, 2015). Madsen (2015) further adds that an interesting topic for investigation is the role of “management fads and fashions” in adopting the BSC in SMEs.
- Implementation: Madsen (2015) argues that future research should focus on how the BSC is implemented in SMEs, accompanied by a typology of BSCs used in SMEs. Another topic focusses on what knowledge of the original BSC (proposed by Kaplan & Norton) is used by SMEs and to what degree SMEs make the BSC fit to their organization. Madsen (2015) further argues that empirical data on the usage of the BSC in SMEs should be collected by future researchers.
- Experiences: According to Madsen (2015) the future research should focus on the pointing out the differences in the perceived benefits between SMEs and large firms. Therefore more success stories of the BSC being used by SMEs is needed which can be retrieved by qualitative interviews (Madsen, 2015).

Madsen (2015) concludes the literature review by stating that the usage of BSC in SMEs is a rather understudied topic which needs more attention. Madsen (2015) further recommends practitioners in the field of SMEs to create awareness and raise knowledge about the BSC in order to potentially reduce the very high failure rate among SMEs.

Looking back at sub-section 4.1.2 it is interesting to see that the found clusters in the network analysis do share similarities with the themes which were defined by Madsen (2015). Cluster 1 and 3 (see table 16) are even exactly found in the thematic analysis performed by Madsen (2015). This shows that the possible **citation bias** within the included studies can affect the eventual outcome of a literature review. Since the reviewer found these themes among the formed clusters meaning that the interpretation of Madsen (2015) of “Implementation” is based on studies which have a partially shared knowledge base.

## 4.2 The repeat review

### 4.2.1 The need for a systematic literature review

As this is a repeat of the case study Madsen (2015) already provided a need for a literature review and by stating the shortcoming of a narrative literature review it also becomes clear that a systematic literature review is needed. In sub-section 4.1.1 it is already stated why the review needed to be conducted namely, because the BSC research is primarily focussed on large firms while the usage of the SMEs is little to non-existent (Madsen, 2015). This is a research gap which hinders researchers and practitioners in the field to understand how to use the BSC in SMEs properly.

### 4.2.2 Defining a scope and research question accompanied by boundaries

The scope of the case study is to first explore the existing literature about the usage of the BSC in SMEs to identify underexplored areas within the literature found (Madsen, 2015). Secondly a research agenda on the underexplored areas is developed to identify and address the need and direction for future research (Madsen, 2015). This aim, for the purposes of the repeat review is translated into the following research question:

“What is the need and direction for future research within the underexplored areas in the field of the usage of the BSC in SMEs?”

The boundaries accompanied by this research question:

- No literature can be used which was finally submitted after the 13<sup>th</sup> of October 2015 (since this would post-date the case study of Madsen) to set the repeat review in the same timeframe as the case study.
- The repeat review is carried out by one reviewer in a timeframe of three months.

#### 4.2.3 Developing a review protocol with in-/exclusion criteria

When looking at the research question (see sub-section 4.2.2) it can be said to be a more broad research question as it is not subjected to a specific intervention or populations of some sorts. Rather the research question aims at exploring the future research field of the BSC in SMEs. This means that the in-/exclusion criteria should not focus too much on finding the right study design or fulfilling certain conditions like interventions or populations tested. Rather the results found by the included study should be synthesized, while taking the shortcomings of the study into account. The provided PICOC criteria (see section 3.2) are described for the repeat review in table 17.

PICOC criteria	Inclusion criteria	Exclusion criteria
Population	-All literature concerning the usage of the BSC in SMEs	-Literature only concerning the usage of the BSC only in large firms -Literature only concerning other types of performance measurement -Literature concerning only the SMEs and not BSC
Intervention	-The usage of the BSC within SMEs	-The usage of the BSC only within large firms -The usage of only other types of performance measurement within SMEs or large firms
Comparison	No comparison, only seeking to identify research gaps	-Comparison with practices of the BSC within large firms
Outcomes	-Identifications of areas within the field of BSC in SMEs -Identifications of research gaps within the field of BSC in SMEs -Developing a future research agenda for the BSC in SMEs	-Usage of only other types of performance measurement practices
Context	-Timeframe for included studies is 13 <sup>th</sup> of October 2015 and before (see sub-section 4.2.2) -The research domain is concerned with business, management or accounting	-Other irrelevant research domains such as biology, psychology, etc.

Table 17: PICOC criteria

Looking at the exclusion criteria in table 17, there can be said that most exclusion criteria are aimed at excluding literature that only concerns other types of performance measurement systems rather than the BSC. Furthermore literature that only concerns the BSC usage in large firms is excluded as the aim of the research is focussed on SMEs.

#### 4.2.4 Identifying and selecting research

As earlier mentioned the main themes in the research question are focused on the usage of the BSC in SMEs. BSC and SMEs are therefore the core concepts on which the search for studies should focus. Table 18 shows the core concepts with adjacent search terms are used to find as much literature for these two concepts. These search terms are based on synonyms or the broader topic under which these concepts reside.

<b>BSC (Balanced Scorecard)</b>	<b>SMEs (Small and Medium Enterprises)</b>
Performance measurement AND scorecard NOT "enterprise and risk management"	small OR medium AND firm OR firms
Performance measurement AND scorecard NOT "Business intelligence"	small OR medium AND company OR companies
Performance measurement AND scorecard NOT "Business analytics"	small OR medium AND organisation OR organisations
	small OR medium AND organization OR organizations
	small OR medium AND business OR businesses
	small OR medium AND factory OR factories
	small OR medium AND corporation OR corporations

**Table 18: Search terms (Boolean Logic)**

Combining the concepts together, with Boolean Logic, will provide a search term like (for example) this: ((Performance measurement AND scorecard) (NOT "enterprise and risk management" AND ((small OR medium) AND (firms OR firms))). All the possible combinations hidden within table 18 are used to gain as many studies from Scopus and Web of Science, which are then subjected to a funnel for filtering out irrelevant studies. The choices for omitting studies in the funnel are clarified in figure 8, accompanied by the number of studies excluded and included.

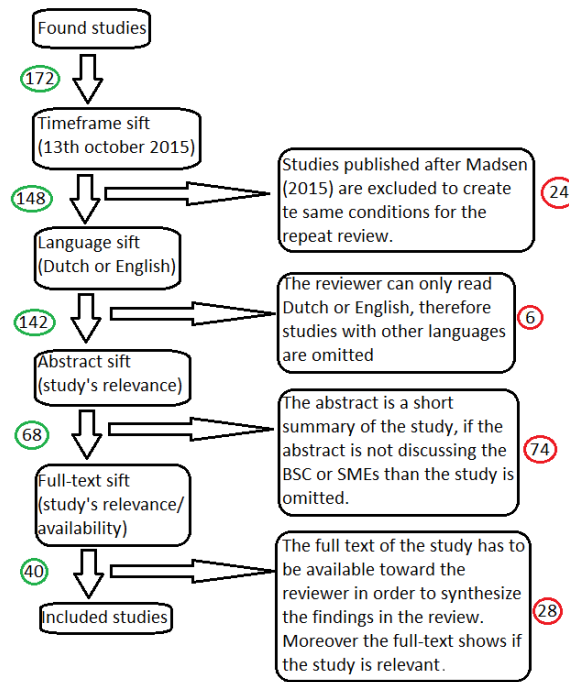


Figure 8: In-/Exclusion criteria

While the first two sifts (see figure 8) are understandable it does perhaps exclude relevant studies as studies in other languages might be relevant as well (this leads to a total of 30 studies being excluded). This poses a source of **language bias** for the repeat review, which affects 3,5% (6 out of 172) of the potential studies for inclusion. The third sift is focussed on excluding studies which aren't relevant by analysing the abstracts which give insight into the focus and results of a study. 74 studies were excluded during this sift mostly because of the way Web of Science works. Web of Science does not only specify the authors key-words but also their own key-words (named: key-words PLUS). Key-words-PLUS often added Balanced Scorecard as a key-word while in conjunction with performance measurement, while the latter is discussed in these studies the BSC is not. Therefore many studies are excluded as they do not discuss the BSC in relation with SMEs. The last sift is a final check to see if the studies full content and results are relevant toward our repeat review. It also clarifies on a study's shortcomings which can influence the way in which its results are represented in the review. It is however necessary to have access to the full text, if this is not the case the study will have to be omitted because the results of it might be untrustworthy. This led to another 28 studies being omitted from the repeat review. After the final sift 40 studies are eligible and usable for the repeat review. 11 of the 40 included studies are also included by the case study of Madsen (2015) (see appendix C) which is discussed more intensively during the comparison in sub-section 4.3.1. As mentioned in section 3.2 there is no secondary searching within the references of the 40 studies which were found due to time constraints of the reviewer.

#### 4.2.5 Assessing the quality of the studies found

Out of the 40 includes studies 30 were peer-reviewed articles which were published in journals with an impact factor. Looking at the average cumulative impact factor (1,527, see appendix B) it can be considered quite high as compared to the average impact factor among journals within the subject area. Among the 1394 journals the average cumulative impact factor of both reviews rank on the

130<sup>th</sup> place<sup>12</sup>, which puts it the top 10%. This poses a source of **place of publication bias** as studies with less significant results are less likely to show up in popular journals (Song et al., 2010).

However when looking at the impact factors of the individual journals it is clear that only 8 out of the 40 included studies in the repeat review have an impact factor above 1,500 (see appendix B). Journals such as Accounting & Finance research have such a high impact factor (> 4,500) that it skews the average while only one of the included studies in the repeat review came from this journal. This in turn decreases the impact of **place of publication bias** for both the repeat review and the case study equally. Furthermore when looking to impact factors in general, 67% of all journals rank above 1,000 according to Journal Citation Report (Web of Science) lessening even more the effect **place of publication bias**.<sup>13</sup> Looking at the individual journals there is no journal which contributed significantly more studies to the repeat review.

#### 4.2.6 Extracting and synthesizing the data

This sub-section first discusses the authors of the included studies to find a potential source of bias. Secondly the age of the included studies will be discussed and thirdly the type of data collection. At last the network analysis and found clusters of the repeat review are discussed.

##### Authors of the included studies

When looking at the authors of the included studies there are only two authors which have provided multiple contributions towards the review. Sousa & Aspinwall worked together on two studies which are included in the repeat review. The literature review of Sousa & Aspinwall (2010) does refer to Sousa et al. (2005) but is not based on the results of Sousa et al. (2005) and uses over 50 different sources to gather data from. This makes it highly unlikely that there is a source of bias coming from this case.

##### Year of publication of the included studies

When looking at the included studies in the repeat review, it does include studies of which some of them are quite old, while others are as recent as the year Madsen (2015) was published (see figure 9).

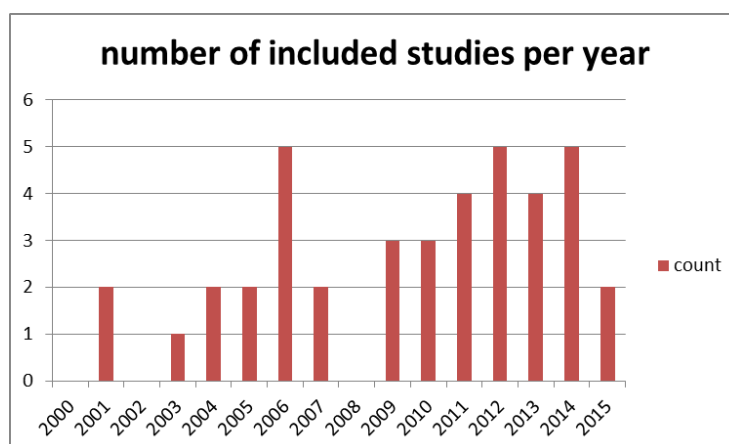


Figure 9: Year of publication of the included studies in the repeat review

<sup>12</sup> Retrieved from [http://www.scimagojr.com/journalrank.php?area=1400&page=3&total\\_size=1394](http://www.scimagojr.com/journalrank.php?area=1400&page=3&total_size=1394) at 28-07-2017

<sup>13</sup> Retrieved from <http://mdanderson.libanswers.com/faq/26159> at 28-07-2017

Looking at figure 9 it becomes clear that there is a quite even distribution among studies which are over 5 years old (2009 and below) at the time of the review of Madsen (2015) and studies under five years old (2010 and up). Therefore the same argument of Pautasso (2013) counts for the repeat review, as this mix of older and more recent studies provide a more complete view of the research area.

#### Type of data collection of the included studies

The studies included in the repeat review show a great variety in which the data was gathered for reaching their results. The various types of data collection of these studies are classified in figure 10 (in the same manner as for the case study, see sub-section 4.1.2).

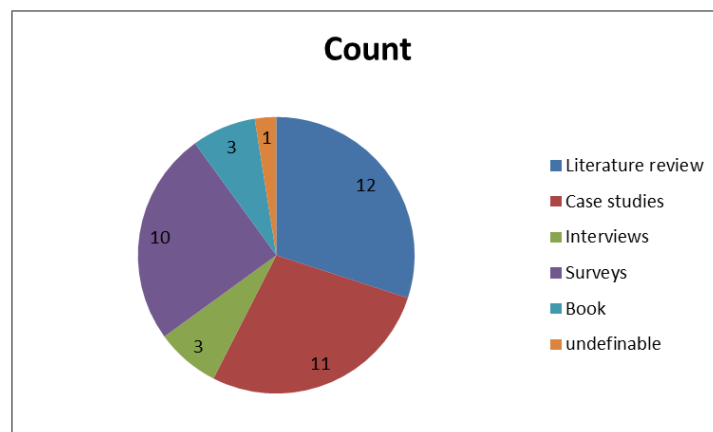


Figure 10: type of data collection of the included studies in the repeat review

When looking at figure 10 the “undefinable” stands for the particular case of Perrini & Tencati (2006) which means that there is a slight **dissemination bias** as the results are not understood in its proper context. However provided that the results are understood out of context and since this is a single case (out of 40) the impact of **dissemination bias** can be considered low.

Looking at figure 6 (see sub-section 4.1.2), from a positivistic stance most of the evidence of the included studies belongs to the second category within conducted cohort studies or case-control studies, because not one of the included studies uses randomized control trials of any kind. According to Evans (2002) case studies, interviews and surveys are all considered poor quality evidence. However Evans (2002) is relating this to the medical field where randomized control trials are more appropriate and easier to design than in the field of business research. Similar to the case study better quality evidence might not be available, even when searching through the two biggest scientific search engines.

#### Network visualisation and cluster analysis

VOS viewer requires an output from either Web of Science or Scopus which are provided for all studies in the repeat review.

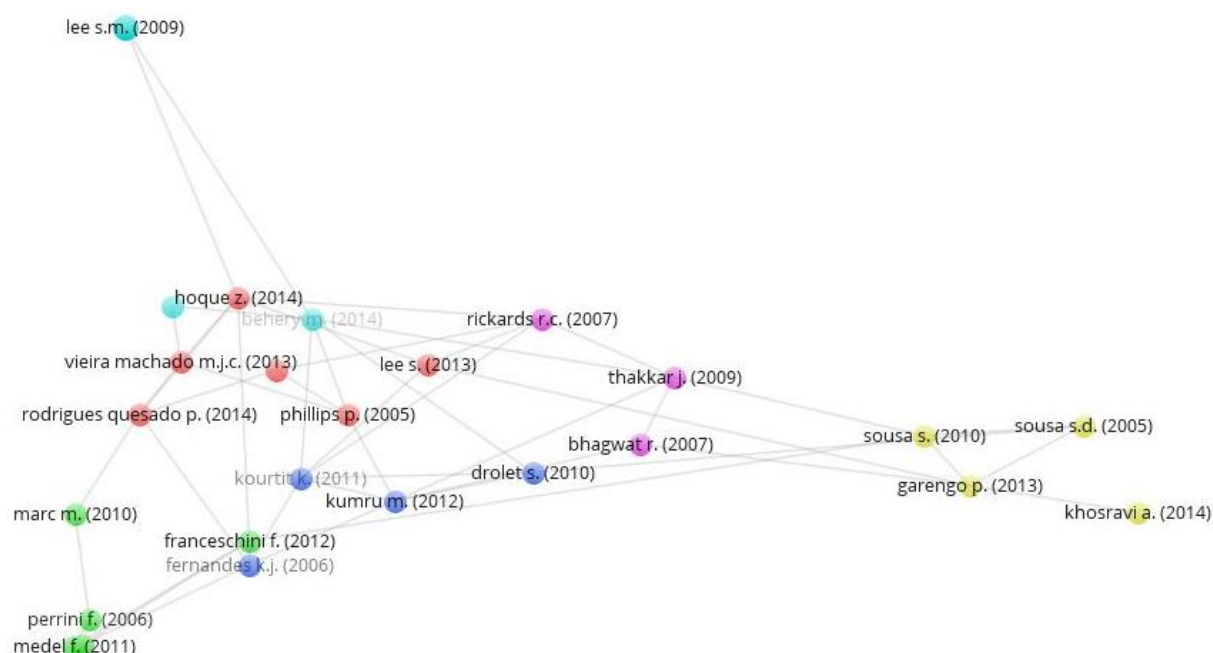


Figure 11: Network visualisation of the repeat review

As can be seen in figure 11 six clusters were formed spanning a total of 26 out of the 40 articles. 14 of the included articles did not form a connection or cluster with one of the other included studies meaning that they don't share a common foundation. The studies within the clusters do share some common foundation and are described in table 19.

Cluster 1: implementation	Cluster 2: Modification	Cluster 3: issues for adoption	Cluster 4: Awareness	Cluster 5: Supply chain BSC	Cluster 6: undefinable
Hoque (2014)	Marc, Peljhan, Ponikvar, Tekavcic & Sobota (2010)	Drolets & LeBel (2010)	Khosravi, , Ghapanchi, Najaftorkaman & Ismail (2014)	Bhagwat & Sharma (2007)	Behery, Jabeen & Parakandi (2014)
Laitinen & Chong (2006)	Franceschini & turina (2012)	Fernandes et al. (2006)	Garengo & Biazzo (2013)	Rickards (2007)	Gumbus & Lussier (2006)
Lee, Park & Lim(2013)	Medel, Caro Garcia & Anido (2011)	Kourtit & Nijkamp (2011)	Sousa & Aspinwall (2010)	Thakkar, Deshmukh & Kanda (2009)	Lee, Kim, Choi & Lee (2009)
Philips & Louvieris (2005)	Perrini & Tencati (2006)	Kumru (2012)	Sousa, Aspinwall & Sampaio(2005)		
Rodrigues Quesado, Guzman & Rodrigues(2014)	Van der Woerd & van den Brink (2004)				
Machado (2013)					

Table 19: Cluster description of the repeat review<sup>14</sup>

<sup>14</sup> The cluster names are based on an interpretation of the common theme by the reviewer after reading the included studies of a cluster.

Looking at the evidence from table 19 the repeat review does include significant clusters of studies including 7,5% (3 out of 40 included studies) to 15,0% (6 out of 40 included studies) of the total number. The forming of clusters does not mean that the studies within the clusters are solely founded on the same foundation, however the closeness of the studies in figure 11 indicates to which degree they share a common foundation. In cluster 1, for example, the studies are relatively close to each other meaning that the studies share a relatively great common foundation, whereas the studies in cluster 6 are further apart. Looking at the total visualisation the clusters (except for cluster 4 and one study of cluster 6) are quite close and even intertwined with each other. This means these authors at least share some common references which they use to support their studies. A common relation among all studies within figure 11 is that they use Kaplan & Norton as a reference which is logical as they developed the BSC in Kaplan & Norton (1992).

The studies in cluster 1 were all addressing the implementation of the BSC in SMEs along with issues during this process. Cluster 2 was mostly concerned with different varieties of the BSC in SMEs, by altering the basic BSC as provided by Kaplan & Norton (1992). Cluster 3 was mostly concerned with why SMEs do not often adopt the BSC. Cluster 4 was only concerned with why and to what extent SMEs are not aware of the BSC. Cluster 5 was primarily concerned with the usage of the BSC within the supply-chain structure of SMEs. In cluster 6 no binding element for clustering was found, meaning that although they share common sources these studies used the evidence in a different way as compared to their clustered associates.

#### 4.2.7 Analysing, reporting and discussing the findings

This sub-section discusses the analysis of the 40 included studies for the repeat review. First the information of the 40 studies of the repeat review is coded by using free-line-coding (see section 3.2). Secondly a thematic analysis is performed while using the results of free-line-coding as a guide. At last a research agenda is formed like in the case study of Madsen (2015).

##### Coding of the results in the included studies

In order to find recurring themes in the body of included literature coding was used to identify these themes. During the process of free-line-coding some themes were identified that were so much interrelated that they together formed a bigger more general code, which was then used to form a theme. Table 20 shows which codes were found among the included studies and which themes they formed. This is accompanied by the number of times the codes were found in the included studies to show the importance of a theme as compared to the other themes found.

Theme	Found	Code(s)	Found
Adoption	29	Adoption rates	14
		Financial vs non-financial measures	15
Benefits	12	Benefits	12
Problems	20	Problems	20
Knowledge & Awareness	13	Knowledge	6
		Awareness	7
Implementation	48	Pre-existing performance measurement	9
		Success factors for implementation	28

		Contextual factors for implementation	11
Changing the BSC	17	Adapting the BSC	11
		Integrating the BSC	6
<b>Total</b>	<b>151</b>	<b>Total</b>	<b>151</b>

Table 20: Coding scheme of the repeat review

As can be seen in table 20 a total of six themes were distinguished which are established on 151 codes of the included studies. The most important theme which was found is the “Implementation” which refers to the implementation of the BSC in SMEs. The included studies discussed a lot of success factors for a implementation of the BSC. Even the smallest theme found “Benefits” which refers to the benefits of implementing the BSC in SMEs poses significant coding results of 8% (12 out of 151). All these themes are therefore discussed in the thematic analysis

Looking back at sub-section 4.2.6 it is interesting to see that the found clusters in the network analysis do share similarities with the themes constructed in the repeat review Cluster 1 to 5 (see table 19) are even exactly found in the thematic analysis. This shows that the possible **citation bias** within the included studies can affect the eventual outcome of a literature review. However in the repeat review, the reviewer first named the clusters before performing the free-line-coding. The naming of the clusters therefore affected the reviewer which possibly lead to constructed themes.

### Thematic analysis

When looking at the result from table 20 the themes found are described in a narrative fashion. This description of themes is used to identify gaps in the research with which later on a research agenda is formed.

Knowledge and awareness: Giannopoulos, Holt, Khansalar & Cleanthous (2013) found that only 20% of the SMEs in the UK know about the BSC. Machado (2013) found that only 27% of the SMEs in Portugal know about the BSC. According to Marc et al. (2010) companies which are aware and have knowledge of management tools (such as the BSC) are more likely to implement such practices. 50% of the non-users are not aware and don't have knowledge about the BSC (Marc et al., 2010).

There is a general lack of knowledge about the BSC in SMEs (Fernandes et al., 2006) (Giannopoulos et al., 2013)(Kourtis & Nijkamp, 2011)(Monte & Fontenete, 2012) (Rodrigues Quesado et al., 2014) (Sousa et al., 2005). When looking at the present knowledge of the BSC among SMEs there is a gap of knowledge on the non-financial measures, such as the customer perspective, innovation and learning (Chimwani, Nyamwange & Robert, 2013). Moreover when SMEs were questioned about it, 75% told that financial measures are very important while non-financial measures were less important. (Giannopoulos et al. 2013) (Sousa et al., 2005).

Adoption of the BSC: When looking at the adoption rate of the BSC in general Kumru (2012) and Monte & Fontenete (2012) state that mainly large companies adopt the BSC while SMEs don't do this. For example, only a marginal number of SMEs in Portugal uses the BSC (Machado, 2013)(Rodrigues Quesado et al., 2014) and only 5% of the SMEs in the UK uses a form of the BSC (Giannopoulos et al., 2013). Niven (2015) however argues, out of personal experience as a scorecard consultant, that he has seen successful adoptions of the BSC within SMEs. Moreover Phillips & Louvieris (2005) deem the BSC to be an appropriate tool for SMEs within the tourism industry.

When looking at the adoption of the BSC in SMEs various authors agree that there is a big focus on financial measurements as compared to non-financial measurements (Chimwani et al., 2013)(Rickards, 2007) (Sousa et al., 2005)(von Bergen & Benco, 2004). An actual survey as presented by Monte & Fontenete (2012) revealed that that financial measures were used by 100% of the SMEs within their study while the non-financial measures were only used by 48,5% up to 75%. Old approaches towards performance measurements were primarily aimed at financial measures, while more recent approaches seek to incorporate non-financial measures as well (Kourtis & Nijkamp, 2011)(Rickards, 2007). For the sake of the long-term strategy it is important for SMEs to adopt non-financial measures in their BSC (Basuony, 2014). These non-financial measures have to be balanced with financial measures of the BSC (Lee et al., 2013).

Implementation: When implementing the BSC in SMEs more than often are pre-existing performance management or key-performance indicators present within SMEs (Franceschini & Turina, 2011)(Behery et al., 2014)(Manville, 2006)(Monte & Fontenete, 2012). These pre-existing indicators and management structures are merged with the BSC approach to create the earlier mentioned benefits. In order to do this SMEs need to align these pre-existing indicator, which are often from different work units in the company, with the strategic vision and objectives (Behery et al.).

Strategic vision and objectives have to be defined by a designated person within the SME (da Costa marques, 2012). Frequent changes in this strategy and objectives are to be avoided when implementing the BSC in SMEs (Rompho, 2011). However, van der Woerd & van den Brink (2004) argue that the strategy and objectives should be checked, and if necessary updated, once every year. When operationalizing the strategy and objectives into measures of the BSC the number of measures should be quite low (5-15) for SMEs (Basuony, 2014) (Liu & Jiang, 2010)(Liu & Yang)(von Bergen & Benco, 2004). The used measurements in SMES should focus on the breadth of performance that they measure rather than providing in-depth knowledge (Basuony, 2014) (da Costa Marques, 2012). The organisation must therefore carefully select and measure the processes which lead to improved outcomes for customers and ultimately allows the organisation to reach its objectives (Niven, 2015). Proper education and awareness about the BSC is needed among the personnel of SMEs for a successful implementation (Behery et al., 2014) (Fernandes, Raja & Whalley, 2006) (Garrengo & Biazzo, 2013)(Sousa & Aspinwall, 2010). This also requires commitment from the management team to promote learning of the BSC within the SME (Garrengo & Biazzo, 2013). Furthermore, the management of the SME must reallocate manpower for collecting data for the BSC (Liu & Jiang, 2010).

The internal and external environment also needs to be taken into consideration when implementing the BSC and has to be adjusted accordingly (Behery et al., 2014). The BCS and its measures should be adapted towards the critical success factors of an industry (Gumbus & Lussier, 2006)(Rickards, 2007). For example, Phillips & Louvieris (2005) adjusted the BSC towards the tourism industry by adding different measures such as “budgetary control”. Von Bergen & Benco (2004) supports this argument by stating that companies are should adjust the measures to suit their strategy.

Benefits: Andersen, Cobbold & Lawrie (2001) and Sousa & Aspinwall (2010) describe the benefits of the BSC for SMEs as being able to provide SMEs a clear sense of strategic direction as well as a more deep understanding of their business model providing a balance between short-term and long-term

goals. According to Andersen et al. (2001), Basuony (2014), da Costa Marques (2012) the BSC provides these benefits by formalizing and structuring the strategic vision and objectives of SMEs.

The implementation of the BSC provides SMEs a greater flexibility in reacting towards changes in the external environment (Andersen et al., 2001)( Fernandes et al., 2006). On the other hand stability is also provided for the internal processes of a SMEs by implementing the BSC ( Fernandes et al., 2006). However Manville (2006) argues that the full benefits of the BSC will not be visible immediately but are becoming visible over a time after at least twelve months.

Problems: According to Giannopoulos et al. (2013) SMEs do not perceive the BSC as applicable. SMEs tend to have more flexible and informal management which might not always be compatible with the high degree of formalization and the complexity of the BSC (da Costa Marques, 2012) (Fernandes et al., 2006). Moreover according to Frost (2003) management systems are often designed for operational control rather than strategic control within SMEs. SMEs trying to implement the BSC therefore often have difficulties with defining their critical success factors and adjacent indicators to measure (Phillips & Louvieris, 2005)(Sousa et al., 2005).

The implementation of the BSC in SMEs requires a relatively large amount of manpower and financial resources (Liu & Jiang, 2010)(Marc et al., 2010) (Sousa et al., 2005). Lack of even basic performance data and measures within SMEs makes the usage of performance measurement systems like the BSC impossible (Drolet & Lebel, 2010) (Liu & Jiang, 2010)(Rickards, 2007). Constrained by the available resources many SMEs run into problems when they have to keep collecting data to keep the BSC measures up to date (Liu & Jiang, 2010) The BSC lacks a mechanism for updating the indicators to keep them relevant which is important for a SMEs with rapid changing environments (Hudson, Smart & Bourne, 2001).

Changing the BSC: Many authors use the BSC model in conjunction with other types of performance management model to create a new integrated model which they argue is more suitable for certain goals. The following examples are provided by the included studies:

- Bhagwat & Sharma (2007) and Thakkar et al. (2009) both developed a BSC approach for the supply chain management within SMEs to offer practitioners in these SMEs a balanced view on the performance of their supply chain.
- Garengo & Biazzo (2013) used the BSC when developing an integrated management system for SMEs among to other performance measurement systems.
- Gurel & Sari (2015) used the BSC as a tool for a small human resource company to determine and structure their strategy by using a SWOT (Strength, Weakness, Opportunity and Threat) analysis of the organization.

On the other hand, authors also adjust (or improve) the BSC to make it more suitable towards the contextual factors of SMEs within certain industries (IT business) or countries (China). The following examples are provided by the included studies:

- Khosravi et al. (2014) suggests an improvement of the BSC which is especially useful for the SMEs which is “Knowledge Management”. Knowledge management is used to transfer the right information about indicators of the BSC to the right person at the right time (Khosravi et al., 2014). Lee et al. (2009) established that the usage of knowledge management when

implementing the BSC in small IT companies led to a better process performance which in turn contributed towards a better financial performance.

- Kumru (2012) developed a BSC approach specifically for small media focussed companies by adding an “composite scoring” element to supplement the BSC.
- Liu & Yang (2012) added the “Relationship, Financing and Government support” to the BSC to better address the situation of SMEs in China.
- Medel et al. (2011) added “sustainability” in the form of “sustainability reporting guidelines” to the existing BSC in order for SMEs to improve and communicate their long-term strategy. However Perrini & Tencati (2006) argue that these guidelines are not suitable for SMEs as they are too complex and require a large number of formal processes.
- The “changing of the BSC” has shown different varieties of usage of the BSC. All these varieties were developed for specific purposes, and moreover, specific cases. Future research should focus on the generalizability of these various kinds of the BSC as well as empirically verifying them.

The main conclusion that can be drawn from this literature review is that although SMEs in general know little about the BSC, let alone adopt the BSC, it can be very beneficial for SMEs to use. This review shows not only benefits but also provides potential hindrances for SMEs wanting to adopt the BSC and other considerations which need to be taken into account, like adapting it to the needs of the SME or the environment of the SME. As described during section 4.2 this review is subjected to multiple sources of bias and an evidence base which might not offer a high quality of evidence (see sub-section 4.2.6) but it does provide an overview of the existing literature concerning the research area of the BSC in SMEs.

### **Research agenda**

The research agenda is based on the thematic analysis and provides directions for future research.

- According to Laitinen & Chong (2006) and Liu & Jiang (2010) little research has been performed on the BSC in SMEs. Future research should focus more on analysing the BSC within SMEs instead of large companies since two thirds of the employable people in the UK work in SMEs (Laitinen & Chong, 2006)
- There exists a gap in knowledge about the BSC within SMEs (Giannopoulos et al., 2013)(Machado, 2013). Kourtis & Nijkamp (2011) argues that especially the link between performance management and strategy is unknown to SMEs. However, multiple other areas are mentioned as well. Therefore, future research should focus on analysing the gap in knowledge about the BSC within SMEs into further detail.
- At the “problems” themes many boundaries were presented why SMEs would not adopt and implement the BSC. Future research should focus on describing these boundaries in further detail and empirically verify or reject them.
- According to Chimwani et al. (2013), Rickards (2007), Sousa et al. (2005) and von Bergen & Benco (2004) there is a great focus on the financial measures within SMEs. Future research should focus on the reason why financial measures as important for SMEs in comparison to non-financial measures.
- According to Rodrigues Quesado et al., (2014) there is a big lack of empirical information on the implementation of the BSC within SMEs. Future research should therefore focus on strategies for implementation that have been empirically proven to work (Hoque, 2014).

### 4.3 Comparison of the case study and the repeat review

This section discusses the comparison between the case study of Madsen (2015) and the repeat review. First both reviews are compared on their selection processes and outcomes. Secondly the interpretation of the information from the included studies in both reviews is compared. This section provides the answer to sub-research question 3 and 4 (see section 1.4).

#### 4.3.1 Comparison of both selection processes and outcomes

##### Analysis of studies included by both the case study and the repeat review

As mentioned in sub-section 4.2.4 the case study and the repeat review share 11 of their included studies. This means that both reviews are partially based on the same base of evidence (see figure 12).

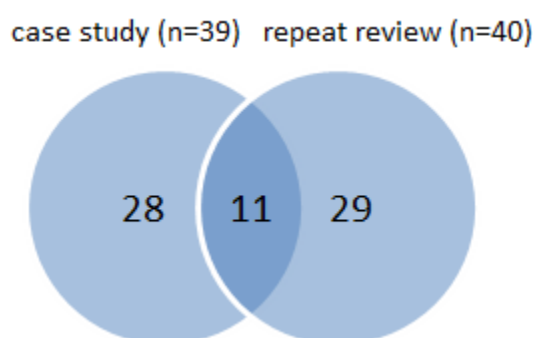


Figure 12: Studies included in both the case study and the repeat review

To find out why both the case study and the repeat review included these studies an analysis is performed on the characteristics of these studies (see table 21, the total set of data can be found in appendix C)

Author(s)	Year of publication	Journal or other source	Type of data collection	Impact factor
No authors were present more than once in the 11 studies.	Year of publication ranges from 2001 to 2014.	No journal or other source was present more than once in the 11 studies. 8 studies were peer reviewed, 3 were not.	5 literature reviews, 2 case studies, 3 interviews and 1 survey were used as the type of data collection.	2 of the studies came from high (>1,000) impact journals. On average the impact was below 1,000.

Table 21: Analysis of the studies included in both reviews

Looking at table 21 there is no pattern visible in the author(s) of the shared studies as they all contributed only once to both reviews. The year of publication also does not show a pattern as the range of years reflect those of both the case study and the repeat review (see figure 4 and 9). No journals were present more than once leading to no visible pattern, also non-peer reviewed studies were included. The distribution of types of data collection in the shared studies pretty much reflects the distribution of types of data collection in both the case study and the repeat review (figure 5 and 10). There is also no presence of studies with only high impact factors. Overall it can be stated that there is no specific reason to be found in the characteristics of these 11 studies that they were chosen for a specific reason. However, since Madsen (2015) did not provide the way in which he

searched for his studies it might be possible that he also searched for studies in Web of Science and Scopus, leading to him finding the same studies.

### Authors of the included studies

Both the case study and the repeat review included multiple studies of the same author(s). However after investigating these multiple contributions it became clear that only the case study shows a problematic case of **multi-publication bias**. Madsen provided 4 studies (partially) of his own writing towards his review which were based on the same set of data, potentially leading to an over-exaggeration of the outcomes in his review.

### Year of publication of the included studies

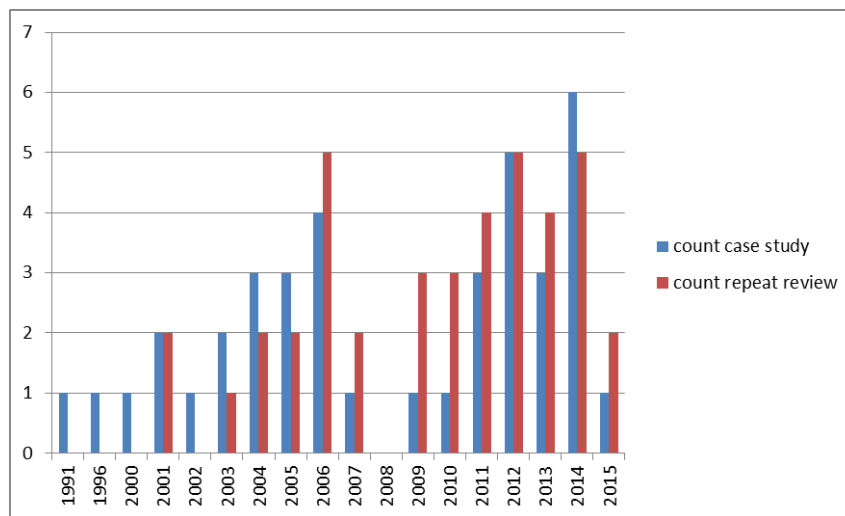


Figure 13: comparison on the year of publication of both reviews

Looking at distribution of the year of publication of both reviews it becomes clear that there only is a slight difference between the case study and the repeat review. The case study includes slightly more older studies than the repeat review, but the average year of publication is only 1,5 year apart between the case study and the repeat review. No **bias** with regard to time can therefore be observed.

### Type of data collection of the included studies

Type of data collection	Repeat review	Case study
Literature review	12	13
Case studies	11	6
Interviews	3	6
Surveys	10	10
Book as a study design?	3	2
Other	1	2
<b>Total included studies</b>	<b>40</b>	<b>39</b>

Table 22: study designs of included studies

Table 22 shows that the biggest difference between the repeat review and case study regarding study designs is that the repeat review relies more on evidence provided by analysing case studies rather than performing interviews. The case study and the repeat review show no further significant



## Network visualisation and cluster analysis

Case study	Repeat review
Implementation (10 studies)	Implementation (6 studies)
Future research (6 studies)	Modification (5 studies)
Adoption (3 studies)	Issues for adopting (4 studies)
	Awareness (4 studies)
	Supply chain BSC (3 studies)
	Undefinable (3 studies)

Table 23: Comparison of the found clusters in both reviews

Looking at table 23 the case study of Madsen (2015) has less clusters than the repeat review, however the first is significantly larger than all the others in both reviews. Both the case study and the repeat review showed clusters with studies that are related to a specific theme in the thematic analysis (except for “Future research” in the case study and “undefinable” in the repeat review). Two of the cluster names are even found in both reviews further establishing the role of the shared included studies. For example the cluster “implementation” includes 3 shared studies, but this also means that the rest of the studies in this cluster are not shared. So even when using partially different sources it still can lead towards the same common theme being found in a cluster.

Although both reviews show a different number and size of clusters it is not easily established whether one is more influenced by **citation bias** than the other because of this. The impact of **citation bias** depends on the role that the shared references play within the studies. For example, Kaplan & Norton is often used to introduce the BSC, rather than to use it as an argument to support a hypothesis like the study of Rodrigues Quesado et al. (2014), Marc et al. (2010) and others. In this case there is no **citation bias** to be observed, hence other shared references can play a role in supporting arguments within studies which does induce **citation bias**. This all means that both reviews are affected by **citation bias** but the impact of it for each review needs to be investigated further.

### 4.3.2 Comparison of both interpretation processes and outcomes

#### Similarities

Before looking at the differences between the two analyses first the commonalities are discussed.

- The themes found by both the repeat review and the case study of Madsen (2015) are very similar. “knowledge & awareness”, “adoption” and “implementation” are themes found in both reviews. While Madsen uses the theme “experiences” he does split these up between “benefits” and “problems” which are also found in the repeat review.
- Both reviews state that awareness and knowledge about the BSC is low among SMEs.
- Both reviews state that adoption rates of the BSC are low among SMEs.
- Both reviews state that when implementing the BSC in SMEs the circumstances (like size of the organisation, knowledge about the BSC among employees, etc.) should be taken into consideration. It should be taken into consideration that the circumstances were not fully specified in the case study of Madsen (2015), thus the repeat review might refer to different circumstances. This poses some potential **interpretation bias**.
- Both reviews state that a benefit of implementing the BSC in SMEs is that the BSC supports the process of setting objects and form a strategy.

- Both reviews state that a lack of knowledge on implementing the BSC among SMEs is a barrier.

As can be seen there is quite a large number of similarities to be found between both reviews. This can partially be explained by the fact that both reviews share 11 of their included studies (see sub-section 4.6.1) thus the thematic analyses of both review are partially based on a common foundation.

### Differences

The following differences can be seen between the repeat review and the case study of Madsen (2015).

- The repeat review further specifies the lack of knowledge about the BSC among SMEs by stating that SMEs primarily focus on financial measures and have little knowledge about non-financial measures of the BSC.
- The repeat review does specify that SMEs do adopt financial measures but don't adopt the other non-financial measures of the BSC.
- While the case study states the importance of internal and external factors at the "adoption" stage of the BSC among SMEs, the repeat review views these as factors which should be taken into account when "implementation" the BSC in SMEs. The case study views these factors more as a deal-maker or deal-breaker, while the repeat review views these factors as an influence on the implementation of the BSC. It should be noted that the internal or external factors were not specified in the case study and specified differently among the studies found in the repeat review. This is a potential source of **interpretation bias** as internal and external factors are different among included studies or not specified into detail.
- The case study states that the "implementation" of the BSC tends to go faster in SMEs and that SMEs are using a simplified version of the BSC. The repeat review however, focuses on explaining critical success factors for implementing the BSC as well as arguing that pre-existing measures are present and should be merged into the BSC.
- While the case study stresses that a main benefit of implementing the BSC for SMEs is that they can avoid the pitfalls of large companies, the repeat review focusses on the flexibility towards a changing environment and the internal stability of the SME when using the BSC.
- While the case study highlights the dynamic environment (can change quickly) as a problem for the usage of the BSC in SMEs, the repeat review found that the informal management of a SME is often not suitable for using the more formalized and sometimes complex BSC.

As can be seen, plenty of differences can be found as well between the repeat review and the case study of Madsen (2015). **This shows that even when both review share a similar aim and goal, the outcome can be different.** The different outcome can be explained as the included studies in both reviews are not fully the same. This is the result of searching process which must have been different between the case study and the repeat review. For example, Madsen (2015) could have used different search terms and sources for obtaining the studies he included in his review.

In contrast to this both reviews do share 11 of their included studies but the way in which the information out of these studies is used sometimes differs between the repeat review and the case

study. For example, both reviews use the article of Giannopoulos et al. (2013) in the thematic analysis. Madsen (2015) refers to Giannopoulos et al. (2013) in the following phrases:

- *“In a survey of companies in the UK and Cyprus, Giannopoulos et al. (2013) found that the companies had little knowledge and awareness about the BSC.”(Madsen, 2015, p.3)*
- *“However, the study by Giannopoulos et al. (2013) found that one main reason for non-adoption is that the BSC is perceived as not applicable for small companies.” (Madsen, 2015, p.3)*
- *“[...]while in the study by Giannopoulos et al. (2013) it was discovered that the majority did not use the BSC.” (Madsen, 2015, p.3)*

These three references by Madsen (2015) to Giannopoulos et al. (2013) refer to the same information which is also included in the repeat review. However the repeat review also refers to Giannopoulos et al. (2013) which states that SMEs find financial measures more important than non-financial measures. Madsen (2015) chose not to include this information in his review while the repeat review deemed it to be important. This shows a case of **interpretation bias** as both reviewers had accessed the same study but interpreted the importance of the information in a different way, which led to the inclusion of extra information in the repeat review. A similar case can be made for the shared included study of Fernandes et al. (2006) and the other shared studies. The impact of **interpretation bias** found in thematic analysis of both reviews can be considered small. Both reviews interpreted the information in the shared included studies in a similar direction, but sometimes chose to include extra information if it regarded a code or topic which was mentioned by other included studies.

## 5 Conclusions, discussion, shortcomings and recommendations

This chapter discusses the answers to research question and its sub-questions whilst putting the answers into context by providing shortcomings for which recommendations are given. First the answers to the sub-questions are given, followed by the answering of the main research question. Secondly the contribution of this research project towards the research field is discussed as well as its contribution to practitioners. Thirdly the shortcomings of this research project are discussed which are split up into shortcomings with regard to the repeat review and shortcomings with regard to the research project in general. At last recommendations are provided for researchers on how to deal with the shortcomings and to give them a direction for future research.

### 5.1 Conclusions and discussion

#### 5.1.1 Answering the sub-questions

First the sub-questions of the main research question are answered to provide a more complete explanation of the answer on the main research question.

##### ***“What types of non-systematic and systematic literature review methods are there?”***

When looking at the typologies for different literature reviews (see also table 1) there is not a general consensus on which types of literature reviews do exist. However these authors do agree that a systematic literature review is one of the types of literature reviews and can be defined as follows: *“A systematic review attempts to collate all empirical evidence that fits pre-specified eligibility criteria in order to answer a specific research question. It uses explicit, systematic methods that are selected with a view to minimizing bias, thus providing more reliable findings from which conclusions can be drawn and decisions made.”* (Green et al., 2008, p. 13). Various other types of systematic literature reviews (like umbrella review, rapid review, etc.) are derived from this definition. This is not the case for non-systematic literature reviews which can be differentiated from systematic literature reviews as follows: *“A non-systematic literature review differs from a systematic literature review in that it is not obligated to be explicit about the methods that are used. It therefore is harder to identify sources of bias within non-systematic literature review which is further enhanced by the lack of formal quality appraisal.”* Literature reviewing methods found which lacked these qualities of systematic reviews (such as a critical review, a scoping review, etc.) were deemed as non-systematic.

##### ***“How does bias influence systematic and non-systematic literature reviews?”***

**Publication and associated biases** (Song et al., 2010) are influencing the selection process of studies for literature reviews. These biases refer to studies which are only published when its effects are significant and in the desired direction, thus omitting less significant studies (or in the wrong direction) creating an over-exaggerated image of the effect and a skewed literature review. When interpreting the information of the included studies only one type of bias becomes apparent which is **interpretation bias**. **Interpretation bias** refers to researchers (or reviewers) ability to synthesize, judge and weigh the results found in a study. Two researchers of different backgrounds might look at the same result in a different way thus drawing different conclusions based on their own background (MacCoun, 1998).

When looking at how bias affects systematic literature reviews as compared to non-systematic literature reviews there is one major difference. Systematic literature reviews are subjected to bias in

the same way as non-systematic literature reviews (Massaro et al., 2016), however systematic literature reviews are transparent in their methods making potential sources of bias easier to identify and possibly deal with when reproducing the review. Non-systematic literature reviews do not have the obligation of being transparent about the methods used for selecting studies making the potential sources of bias possibly untraceable. Furthermore it makes reproducing the non-systematic literature review very hard as the steps of the original are not fully specified.

***“What differences do occur when selecting studies for the literature review between a non-systematic literature review and a systematic literature review?”***

When comparing the case study (non-systematic literature review) to the repeat review (systematic literature review), multiple differences were found, as well as similarities. Since the case study of Madsen (2015) did not specify a search strategy the included studies by Madsen (2015) are compared to the included studies in the repeat review which led to the following differences being found:

- The repeat review excluded 3,5% of the potential studies on the basis of language (**language bias**). The effect of this bias could not be measured in the case study.
- Both the repeat review and the case study included peer reviewed studies which were of relatively high ranking journals which roughly affected both review equally (8 and 7 studies respectively)(**place of publication bias**).
- The type of data collection of one study of the included studies in the repeat review was not defined (**dissemination bias**).
- The case study included 4 studies which were based on the same survey data (**multiple publication bias**). Furthermore these for studies were (co-)written by Madsen himself.
- The included studies in both the repeat review and the case study showed clusters of related studies varying in number of clusters and size of clusters. These cluster did provide a base for the thematic analysis of both reviews(**citation bias**)

***“What differences do occur when interpreting the results from the selected studies material a non-systematic literature review and a systematic literature review?”***

When looking at the interpretation of the included studies in the thematic analyses of both reviews all the themes from the case study were also formed in the repeat review (plus one new category). Since both reviews share 11 of their included studies it is logical that similar themes were formed. Another factor adding to this is the fact that the maker of the repeat review also analysed the case study of Madsen (2015).

However many differences between both reviews were also present, even among the interpretation of information in the shared included studies. While the case study took a specific set of information from one of the shared studies, the repeat review deemed other extra information as important to incorporate as well (**interpretation bias**). During this thesis projected only the contribution of the shared studies could be compared side by side. The non-shared included studies by both reviews are also prone to **interpretation bias** as another reviewer might use other information from the non-shared studies as well.

### 5.1.2 Answering the main research question

After answering the sub-questions now the main research question can be answered.

***“What is the impact of bias when selecting studies for a literature review as well as interpreting the results from these selected studies for non-systematic literature reviews as compared to systematic literature reviews?”***

Taken into account that this is a one **case-by-case comparison** the following can be said. When looking at the reasoning from a theoretical point of view, authors argue that systematic literature reviews are subjected to bias as well as non-systematic literature review. However, since systematic literature reviews are required to specify every step of the reviewing process, sources of bias become detectable which is not the case for non-systematic literature reviews as they are not required to specify each step taken. The evidence from the analysis of the case study showed that when selecting and interpreting studies the impact of bias for the case study was as follows.

- 7 of the included studies are providing a potential source of **place of publication bias**.
- 4 of the included studies are a source of **multi-publication bias**.
- 3 clusters of included studies (19 total studies) were formed providing a source of **citation bias** of which the impact remains unclear.
- **Interpretation bias** does affect the at least 11 of the included studies but the other 28 included studies might also be affected as well.

The evidence from the analysis of the repeat review showed that when selecting and interpreting studies the impact of bias for the repeat review was as follows.

- 6 of the potential studies for inclusion were excluded because of their language which poses a source of **language bias**<sup>17</sup>.
- 8 of the included studies are providing a potential source of **place of publication bias**.
- 1 of the included studies provides a very small case of **dissemination bias**.
- 6 clusters of included studies (25 total studies) were formed providing a source of **citation bias** of which the impact remains unclear.
- **Interpretation bias** does affect the at least 11 of the included studies but the other 29 included studies might also be affected as well.

When comparing the impact from both reviews it becomes clear that some types of bias (like **place of publication bias**, **citation bias** and **interpretation bias**) have a similar impact. However, other types of bias (like **language bias**, **multiple publication bias** and **dissemination bias**) have a different or even no impact. This defies the argument made by authors like Tranfield et al. (2003), Petticrew & Roberts (2006) and Mulrow (1994) that non-systematic literature reviews are inferior to systematic literature reviews. This **one case-by-case comparison** showed results concluding to: even when using the rigour of systematic literature reviews bias still occurs, however systematic literature reviews are more transparent on the methods being used, making bias easier to trace and to deal with.

---

<sup>17</sup> Could not be measured in the case study as the search strategy was not provided by Madsen (2015)

### 5.1.3 Theoretical and practical relevance

One of the first theoretical contributions of this thesis project is collecting and synthesizing many different types of literature reviews that provided a typology accompanied with a structured description (SALSA) of all literature reviews found. The second theoretical contribution provided by this thesis project is that it shows the application of a systematic literature review following the steps as proposed by various authors. The third theoretical contribution made by this thesis project is that it adds to the discussion whether there is a need for more rigour in literature review that systematic literature reviews supply, by stating that based on this one case study there is no need for more rigour literature reviews.

Since this is a mostly theoretical thesis project the practical contributions are smaller, but still present. The first practical contribution is provided by the repeat review which adds to the review of Madsen (2015) by adding that practitioners within SMEs should also focus on non-financial measures instead of only focussing on financial measures. The second contribution is also provided by the repeat review in that it states that the BSC can be used and changed towards a specific goal or context by practitioners in the field.

### 5.2 Shortcomings

When looking at the shortcomings they can be divided into shortcomings with respect to the repeat review and with respect to the thesis project in general. The repeat review was as mentioned earlier a rapid literature review which compromises elements of a systematic literature review, because of resource restrictions (see sub-section 2.3.3). The following compromises/shortcomings affected the repeat review.

- When finding the studies only Scopus and Web of Science were used, while other databases could find even more studies for inclusion.
- Only studies which provided full-text access have been included, thus omitting potentially relevant studies.
- The search terms used for finding the relevant studies were only based on lists of synonyms while thesaurus sheets could provide even more relevant search strings.
- Sifting through reference lists of primary studies found in Scopus and Web of Science has not been performed which could have provided more relevant studies (even though it might lead to more **citation bias**).
- Formal quality appraisal has not been performed which could have led to certain studies, which are included in the repeat review being omitted, because they were of inferior quality. This in turn could influence the outcome of the repeat review.
- Since the maker of the repeat review first analysed the case study of Madsen (2015) the maker of the repeat review is influenced when interpreting the information of the included studies.
- Due to the inexperience of the maker of the repeat review certain common topics within the included studies could have been overlooked during the coding process.

With regard to thesis project the following compromises/shortcomings affected it.

- Certain types of bias (like **publication bias**, **funding bias**, etc.) were not measured when comparing the repeat review to the case study since this would require an entire new research project.
- Using network visualisations has provided some interesting results, but the actual impact of **citation bias** could not be given
- Since this research project only presents one case-by-case comparison the generalizability of the found results are quite low.
- The impact of **interpretation bias** could not be measured among both reviews.
- This was the first time this reviewer performed analysis on an existing review and carried one out himself, thus the reviewer lacked experience which might have led to him overlooking or omitting certain aspects (like quality appraisal) of literature reviewing.

### 5.3 Recommendations

Like the shortcomings the recommendations can also be split up between the repeat review and the thesis project in general. It is recommended to re-do the repeat review but in a fully systematic way not compromising on certain elements. This will make the differences between systematic and non-systematic literature reviews more clear.

- Add more sources than only Scopus and Web of Science.
- Add a thesaurus for the term BSC and SMEs.
- Include studies found by searching in the reference lists of primary studies.
- A form of quality appraisal specific for business studies needs to be added to omit studies of inferior quality.
- The future researcher which repeats the repeat review should only be given the aim, research question and boundaries of the original study by Madsen (2015) in order to prevent the influence of the original study on the repeat review.
- The Coding scheme should be checked by multiple scholars to prevent overlooking certain codes.

With regard to thesis project the following recommendations can be given.

- The types of bias affecting literature reviews which were not measured should be measured as well to create a more complete image on the impact of bias on systematic versus non-systematic literature reviews. A meta-analysis for example can be used to assess **publication bias** among others (Ahmed, Sutton & Riley, 2012).
- Network visualisations can send a reviewer in a direction of potential sources of **citation bias** but after the clusters are found, the common foundation should be investigated further to really conclude whether there really is **citation bias**.
- **Instead of only one repeat review on a case study, multiple should be performed and all be compared to each other.** This increases the validity of the found results and gives a better representation of the impact of bias on systematic versus non-systematic literature reviews.
- The impact of interpretation bias can be measured when the studies for a literature review have been selected. The same body of included studies should then be synthesised and analysed by multiple unrelated researchers to create a more complete image of the impact of **interpretation bias**.

## Appendices

### Appendix A: Extraction form

Author(s):
Year of Publication
Journal:
Title:
Key words:
Abstract:

## Appendix B: Impact factor analysis

As earlier stated a full quality appraisal of the studies found is not conducted, hence a comparison between the studies found for the repeat review and the studies used by Madsen (2015) is made. When looking at the repeat review 30 studies were peer-reviewed and 10 were not. Madsen (2015) included 29 peer-reviewed articles and 10 non-peer-reviewed. Further analysis of the peer-reviewed studies about the impact factor of the journals in which they are published (as mentioned in section 3.2) yielded the following results<sup>18</sup>

Symbol	Journal	Count	%	Impact factor	Impact cumulative
AFR	Accounting and Finance Research	1	3,33%	4,546	4,546
IJB	Benchmarking: An International Journal	2	6,67%	0,436	0,872
BMR	Business and Management Review	1	3,33%	0,988	0,988
BSE	Business Strategy and the Environment	1	3,33%	2,228	2,228
CIE	Computers & Industrial Engineering	1	3,33%	1,542	1,542
ESD	European Journal of Sustainable Development	1	3,33%	1,425	1,425
FPE	Forest Policy and Economics	1	3,33%	1,06	1,06
IMA	Information & Management	1	3,33%	1,628	1,628
IBM	International Journal of Business and Management	2	6,67%	0,616	1,232
IMS	International Journal of Management Sciences and Business Research	1	3,33%	0,125	0,125
IOP	International Journal of Operations & Production Management	1	3,33%	2,191	2,191
IPP	International Journal of Productivity and Performance Management	2	6,67%	0,607	1,214
JBE	Journal of Business Ethics	1	3,33%	2,687	2,687
JSM	Journal of Small Business Management	1	3,33%	1,684	1,684
JTR	Journal of Travel Research	1	3,33%	3,04	3,04
PPM	Problems and Perspectives in Management	1	3,33%	0,145	0,145
RBM	Review of Business Management	2	6,67%	0,765	1,53
SBE	Small Business Economics	1	3,33%	2,15	2,15
SCH	Strategic Change	1	3,33%	0,145	0,145
TEC	Technovation	1	3,33%	1,557	1,557
BAR	The British Accounting Review	1	3,33%	0,746	0,746
JSI	The Service Industries Journal	1	3,33%	0,599	0,599
TQM	Total Quality Management & Business Excellence	3	10,00%	0,652	1,956
WRM	Water Resources Management	1	3,33%	1,355	1,355
	<b>Total</b>	30	100,00%	1,372	1,527

Table 24: Impact factor analysis of the repeat review

Symbol	Journal	Count	%	Impact factor	Impact cumulative
AMR	Academy of Management Review	1	3,45%	8,041	8,041
AAA	Accounting, Auditing and Accountability Journal	1	3,45%	0,854	0,854
ADS	Administrative Sciences	1	3,45%	0,689	0,689
BMR	Business and Management Review	1	3,45%	0,988	0,988
DMB	Danish Journal of Management & Business	1	3,45%	0,12	0,12
IBM	International Journal of Business and Management	2	6,90%	0,616	1,232
IBP	International Journal of Business Performance Management	2	6,90%	0,215	0,43
IEB	International Journal of Electronic Business Management	1	3,45%	0,154	0,154
IMC	International Journal of Management Concepts and Philosophy	1	3,45%	0,098	0,098
IMR	International Journal of Management Reviews	1	3,45%	3,357	3,357
IPP	International Journal of Productivity and Performance Management	5	17,24%	0,607	3,035
IQR	International Journal of Quality & Reliability Management	1	3,45%	0,359	0,359
JBR	Journal of Business Research	1	3,45%	1,815	1,815
JSB	Journal of Small Business Management	1	3,45%	1,684	1,684
JKM	KM Online Journal of Knowledge Management	1	3,45%	0,304	0,304
MAR	Management Accounting Research	2	6,90%	2,494	4,988
PPM	Problems and Perspectives in Management	2	6,90%	0,145	0,29
RBM	Review of Business Management	1	3,45%	0,765	0,765
STF	Strategic Finance	1	3,45%	0,145	0,145
TEC	Technovation	1	3,45%	1,557	1,557
BAR	The British Accounting Review	1	3,45%	0,746	0,746
	<b>Total</b>	29	100,00%	1,226	1,507

Table 25: Impact factor analysis of the case study (Madsen, 2015)

<sup>18</sup> Impact factors are retrieved from <http://www.scimagojr.com/journalrank.php> at 28-07-2017

### Appendix C: Data sheet of 11 studies included in both reviews

Authors	Year	Journal/other source	Peer reviewed	data collection	Language	Impact factor
Andersen et al.	2001	2GC Conference Paper, Presented at SMESME Copenhagen	No	Literature review	English	N/A
da Costa-Marques	2012	Business and Management Review	Yes	Literature review	English	0,988
Fernandes et al.	2006	Technovation	Yes	Literature review	English	1,557
Giannopoulos et al.	2013	International Journal of Business and Management	Yes	Quantitative surveys	English	0,616
Gumbus & Lussier	2006	Journal of Small Business Management	Yes	Case studies	English	1,684
Hoque	2014	The British Accounting Review	Yes	Literature review	English	0,746
Machado	2013	Review of Business Management	Yes	Qualitative semi-structured interviews	English	0,765
Manville, G.	2006	International Journal of Productivity and Performance Management	Yes	Qualitative semi-structured interviews, narrative analysis (in a case organisation)	English	0,607
Monte & Fontenete	2012	Proceedings of World Academy of Science, Engineering and Technology	No	Case studies and surveys	English	N/A
Rompho	2011	International Journal of Business and Management	Yes	Qualitative semi-structured interviews, narrative analysis (in a case organisation)	English	0,616
Von Bergen & Benco	2004	Proceedings of the United States Association for Small Business and Entrepreneurship Conference	No	Literature review	English	N/A

## References

- Ahmed, I., Sutton, A.J., Riley, R.D. (2012) Assessment of publication bias, selection bias, and unavailable data in meta-analyses using individual participant data: a database survey. *The British Medical Journal*, 344, pp. 1-10
- Arksey, H. and O'Malley, L. (2005) Scoping studies: towards a methodological framework, *International Journal of Social Research Methodology*, 8(1) pp. 19-32
- Babbie, E. (2010) *The Practice of Social Research*. Wadsworth: Cengage Learning.
- Baumeister, R.F., Leary, M.R. (1997) "Writing Narrative Literature Reviews". *Review of General Psychology*, 1(3) pp. 311-320
- Booth, A., Papaioannou, D., Sutton, A. (2012) *Systematic Approaches to a Successful Literature Review*. Sage Publications Inc.
- Borenstein, M., Hedges, L., Higgins, J., Rothstein, H.R. (2009) "Introduction to Meta-Analysis". Wiley, Chichester.
- Cronin, P., Ryan, F., Coughlan, M. (2008) "Undertaking a literature review: a step-by-step approach". *British Journal of Nursing*, 17(1) pp. 38-43
- Croom, S., Romano, P., Giannakis, M. (2000) "Supply chain management: an analytical framework for critical literature review". *European Journal of Purchasing & Supply Management*, 6, pp. 67-83
- Dahlander, L., Gann, D.M. (2010) How open is innovation? *Research Policy* 39 (2010), 699-709
- Dicheva, D., Dichev, C., Agre, G., Angelova, G. (2015) "Gamification in Education: A Systematic Mapping Study". *Educational Technology & Society*, 18(3) pp. 1-14
- Dijkers, M.P.J.M. (2009) "The task force on systematic reviews and guidelines: the value of traditional reviews in the era of systematic reviewing", *Am. J. Phys. Med. Rehabil.*, 88, pp. 423-430.
- Dixon-Woods, M., Booth, A., and Sutton, A.J. (2007) Synthesising qualitative research: a review of published reports. *Qualitative Research*, 7, 375-422.
- Egger, M., Davey-Smith, G. (1998) Meta-analysis: bias in location and selection of studies. *British Medical Journal*, 316, pp. 61-66
- Evans, D. (2002) Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions. *Journal of Clinical Nursing*, 12(1), pp. 77-84
- Fink, A. (2005) "Conducting Research Literature Reviews: From the Internet to Paper". Sage Publications, London
- Grant, M.J., Booth, A. (2009) "A typology of reviews: an analysis of 14 review types and associated methodologies". *Health Information and Libraries Journal*, 26, pp.91-108

Green S, Higgins JPT, Alderson P, Clarke M, Mulrow CD, Oxman AD. (2008) Chapter 1: Introduction. In: Higgins JPT, Green S (editors), *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.0.1 [updated September 2008]. The Cochrane Collaboration, 2008.

Hart, C. (1998) "Doing a Literature Review: Releasing the Social Science Research Imagination". *Sage Publications*, London

Institute for Scientific Information (ISI) (1994) Impact Factor. Current Contents print editions June 20, 1994. Retrieved from: <http://wokinfo.com/essays/impact-factor/> on 06-06-2017

Kaplan, R. S. & D. P. Norton. (1992) The balanced scorecard - Measures that drive performance. *Harvard Business Review* (January-February): 71-79.

Kitchenham, B. (2004) Procedures for Performing Systematic Reviews. Software Engineering Group, Department of Computer Science, Keele University.

Kovacs, A., Van Looy, B., Cassiman, B. (2014) Exploring the scope of open innovation: A bibliometric review of a decade of research. *Draft paper* (January 2014), 1-55

Lexchin, J., A Bero, L., Djulbegovic, B., Otavio, C. (2003) Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *British Medical Journal*, 326, pp. 1-10

MacCoun, R. J. (1998) Biases in the interpretation and use of research results. *Annual Review of Psychology*, 49, pp. 259-287

Madson, D.Ø. (2015) "The Balanced Scorecard in the Context of SMEs: A Literature Review". *Review of Business Research*, 15(3), pp. 75-86

Malidou, A. (2014) "Mapping the landscape of knowledge synthesis". *Nursing Management*, 21(5) pp. 30-39

Marr, B., Schiuma, G. (2003) Business performance measurement – past, present and future. *Management Decision* 41/8 (2003), 680-687

Massaro, M., Dumay, J.C., Guthrie, J. (2016) "On the shoulders of giants: undertaking a structured literature review in accounting". *Accounting Auditing & Accountability Journal*, 29(5) pp. 767-801

Murlow, C.D. (1994) "Systematic Reviews: Rationale for systematic reviews". *Bio-Medical Journal*, 309, pp. 597-599

Neely, A. (2005) "The evolution of performance measurement research: Developments in the last decade and a research agenda for the next", *International Journal of Operations & Production Management*, 25, 12, 1264 – 1277

Nudurupati, S.S., Bititci, U.S., Kumar, V., Chan, F.T.S. (2010) "State of the art literature review on performance measurement". *Computers & Industrial Engineering* xxx (2010) xxx-xxx (in press)

O’Cathain A, Thomas K.J., Drabble S.J., Rudolph, A., Hewison, J. (2013) "What can qualitative research do for randomised controlled trials? A systematic mapping Review". *BMJ Open* 3 (2013) pp. 1-15

- Okoli, C., Schabram, K. (2010) "A Guide to Conducting a Systematic Literature Review of Information Systems Research," . *Sprouts: Working Papers on Information Systems*, 10(26).
- Paré, G., Trudel, M., Jaana, M., Kitsiou, S. (2015) "Synthesizing information systems knowledge: A typology of literature reviews". *Information & Management*, 52, pp. 183–199
- Pautasso (2013) Ten Simple Rules for Writing a Literature Review. *PLOS Computational Biology*, 9(7), pp. 1-4
- Petticrew, M., Roberts, H. (2006) *Systematic Reviews in the Social Sciences: A practical guide*. Blackwell Publishing
- Petticrew, M., Roberts, H. (2008) "Systematic reviews—do they 'work' in informing decision-making around health inequalities?". *Health economics, policy, and law*, 3(2) pp. 197-211
- Rocco, T.S., Plakhotnik, M.S. (2009) "Literature Reviews, Conceptual Frameworks, and Theoretical Frameworks: Terms, Functions and Distinctions". *Human Resource Development Review*, 8(1) pp. 120-130
- Rowe, F. (2014) "What literature review is not: diversity, boundaries and recommendations". *European Journal of Information Systems*, 23(3) pp. 241-255
- Scandura, T.A., Williams, E.A. (2000) "Research Methodology in Management: Current Practices, Trends and Implications for Future Research". *Academy of Management Journal*, 43(6) pp. 1248-1264
- Song, F., Parekh, S., Hooper, L., Loke, Y.K., Ryder, J., Sutton, A.J., Hing, C., Kwok, C.S., Pang, C., Harvey, I. (2010) Dissemination and publication of research findings: an updated review of related biases. *Health Technology Assessment*, 14(8), pp. i-220
- Torraco, R. J. (2005) Writing integrative literature reviews: Guidelines and examples. *Human Resource Development Review*, 4, pp. 356-367
- Tranfield, D., David D., & Palminder, S. (2003) "Towards a methodology for developing evidence-informed management knowledge by means of systematic review." *British journal of management* 14.3 (2003): 207-222.
- Waltman, L., Van Eck, N.J., & Noyons, E.C.M. (2010) A unified approach to mapping and clustering of bibliometric networks. *Journal of Informetrics*, 4(4), 629-635
- Watt, A., Maddern, G., Cameron, A., Sturm, L., Babidge, S., Facey, K., Hailey, D., Nordehaug, I. (2008) "Rapid reviews versus full systematic reviews: an inventory of current methods and practice in health technology assessment". *International Journal of Technology Assessment in Health*, 24(3) 133-139
- Webster, J., Watson, R.T. (2002) "Analyzing the Past to Prepare for the Future: Writing a Literature Review". *MIS Quarterly*, 26(2) (Jun., 2002), pp. 13-23
- Wickremasinghe, D., Kuruvilla, S., Mays, N., Avan, B.I. (2015) "Taking knowledge users' knowledge needs into account in health: an evidence synthesis framework". *Health Policy and Planning*, 31, pp. 527–537

## References of the repeat review

- Andersen, H., Cobbold, I., Lawrie, G. (2001) "Balanced Scorecard implementation in SMEs: reflection on literature and practice". *Proceedings of the Fourth SMESME International Conference*, 4, pp. 103-112
- Basuony, M.A.K. (2014) "The Balanced Scorecard in Large Firms and SMEs: A Critique of the Nature, Value and Application". *Accounting and Finance Research*, 3(2), pp. 14-22
- Behery, M., Jabeen, F., Parakandi, M. (2014) "Adopting a contemporary performance management system: A fast-growth small-to-medium enterprise (FGSME) in the UAE". *International Journal of Productivity and Performance Management*, 63(1), pp. 22-43
- Bhagwat, R., Sharma, M.K. (2007) "Performance measurement of supply chain management: A balanced scorecard approach". *Computers & Industrial Engineering*, 53, pp. 43-62
- Chimwani, P.M., Nyamwange, O., Robert, O. (2013) "Application of Strategic Performance Measures in Small and Medium-Sized Manufacturing Enterprises in Kenya: The Use of the Balanced Scorecard Perspectives". *International Journal of Management Sciences and Business Research*, 2(6), pp. 45-60
- Da Costa Marques, M. (2012) "Strategic Management and Balanced Scorecard: The Particular Case of Small and Medium Enterprises (SMEs) In Portugal". *Business and Management Review*, 2(1), pp. 50 – 62
- Drolet, S., LeBel, L. (2010) "Forest harvesting entrepreneurs, perception of their business status and its influence on performance evaluation". *Forest Policy and Economics*, 12, pp. 287-298
- Fernandes, K.J., Raja, V., Whalley, A. (2006) "Lessons from implementing the balanced scorecard in a small and medium size manufacturing organization". *Technovation*, 26, pp. 623-634
- Franceschini, F., Turina, E. (2011) "Proposal for a performance dashboard for the monitoring of Water and Sewage service Companies (WaSCs)". *Water Resources Management*, pp. 1-23
- Frost, F.A. (2003) "The use of strategic tools by small and medium-sized enterprises: an Australasian study". *Strategic Change*, 12, pp. 49-62
- Garengo, P., Biazzo, S. (2013) "From ISO quality standards to an integrated management system: an implementation process in SME". *Total Quality Management & Business Excellence*, 24(3), pp. 310-335
- Giannopoulos, G., Holt, A., Khansalar, E., Cleanthous, S. (2013) "The Use of the Balanced Scorecard in Small Companies". *International Journal of Business and Management*, 8(14), pp. 1-22
- Gumbus, A., Lussier, R.N. (2006) "Entrepreneurs Use a Balanced Scorecard to Translate Strategy into Performance Measures". *Journal of Small Business Management*, 44(3), pp. 407-425
- Gurel, B., Sari, I.U. (2015) "Strategic Planning for Sustainability in a Start-Up Company: A Case Study on Human Resources Consulting Firm". *European Journal of Sustainable Development*, 4(2), pp. 313-322

- Hoque, Z. (2014) "20 years of studies on the balanced scorecard: Trends, accomplishments, gaps and opportunities for future research". *The British Accounting Review*, 46, pp. 33-59
- Hudson, M., Smart, A., Bourne, M. (2001) "Theory and practice in SME performance measurement systems". *International Journal of Operations & Production Management*, 21(8), pp. 1096-1115
- Khosravi, A., Ismail, M.A.B., Ghapanchi, A.H., Najaftorkaman, M. (2014) "A TAXONOMY OF KNOWLEDGE MANAGEMENT OUTCOMES FOR SMES". *PACIS 2014 Proceedings*, Paper 162.
- Kourtit, K., Nijkamp, P. (2011) "Drivers of Innovation, Entrepreneurship and Regional Dynamics". *Springer-Verlag Berlin Heidelberg*
- Kumru, S. (2012) "A balanced scorecard-based composite measuring approach to assessing the performance of a media outlet". *The Service Industries Journal*, 32(5), pp. 821-843
- Laitinen, E.K., Ghong, G. (2006) "How do Small Companies Measure Their Performance?". *Problems and Perspectives in Management*, 4(3), pp. 49-68
- Lee, S.M., Kim, J., Choi, Y., Lee, S. (2009) "Effects of IT knowledge and media selection on operational performance of small firms". *Small Business Economics*, 32, pp. 241-257
- Lee, S., Park, S.B., Lim, G.G. (2013) "Using balanced scorecards for the evaluation of "Software-as-a-service"". *Information & Management*, 50, pp. 553-561
- Liu, N., Jiang, T. (2010) "Enhance SME's Competitiveness Using BSC". *Orient Academic Forum*
- Liu, J., Yang, J. (2012) "A framework for applying Balanced Scorecard to Chinese SMEs". *Management School of Wuyi University, Jiangmen, Guangdong, China*
- Machado, M.J.C.V. (2013) "Balanced Scorecard: an empirical study of small and medium size enterprises". *Review of Business Management*, 15(46), pp. 129-148
- Manville, G. (2006) "Implementing a balanced scorecard framework in a not for profit SME". *International Journal of Productivity and Performance Management*, 56(2), pp. 162-169
- Marc, M., Peljhan, D., Ponikvar, N., Sobota, A., Tekavcic, M. (2010) "Determinants Of Integrated Performance Measurement Systems Usage: An Empirical Study". *EABR & ETLC Conference Proceedings*, pp. 598-609
- Medel, F., Caro, S.E., Garcia, L., Anido, M. (2011) "Reporting Models for Corporate Sustainability in SMEs". *Information Technologies in Environmental Engineering*, pp. 407-418
- Monte, A.P., Fontenete, C. (2012) "Balanced Scorecard in SMEs – A Proposal for Small Gas Stations in Portugal". *International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering*, 6(6), pp. 1083-1094
- Niven, P.R. (2015) "BALANCED SCORECARD STEP-BY-STEP FOR GOVERNMENT AND NONPROFIT AGENCIES". *John Wiley & Sons, Incorporated*

- Perrini, F., Tencati, A. (2006) "Sustainability and Stakeholder Management: the Need for New Corporate Performance Evaluation and Reporting Systems". *Business Strategy and the Environment*, 15, pp. 296-308
- Phillips, P., Louvieris, P. (2005) "Performance Measurement Systems in Tourism, Hospitality, and Leisure Small Medium-Sized Enterprises: A Balanced Scorecard Perspective". *Journal of Travel Research*, 44, pp. 201-211
- Rickards, R.C. (2007) "BSC and benchmark development for an e-commerce SME". *Benchmarking: An International Journal*, 14(2), pp. 222-250
- Rodrigues Quesado, P., Guzman, B.A., Rodrigues, L.L. (2014) "Determinant Factors of the Implementation of the Balanced Scorecard in Portugal: empirical evidence in public and private organizations". *Review of Business Management*, 16(51), pp. 199-222
- Rompfo, N. (2011) "Why the Balanced Scorecard Fails in SMEs: A Case Study". *International Journal of Business and Management*, 6(11), pp. 39-46
- Sousa, S., Aspinwall, E. (2010) "Development of a performance measurement framework for SMEs". *Total Quality Management & Business Excellence*, 21(5), pp. 475-501
- Sousa, S., Aspinwall, E., Sampaio, P.A., Guimarães Rodrigues, A. (2005) "Performance measures and quality tools in Portuguese small and medium enterprises: survey results". *Total Quality Management & Business Excellence*, 16(2), pp. 277-307
- Thakkar, J., Kanda, A., Deshmukh, S.G. (2009) "Supply chain performance measurement framework for small and medium scale enterprises". *Benchmarking: An International Journal*, 16(5), pp. 702-723
- Van der Woerd, F., van den Brink, T. (2004) "Feasibility of a Responsive Business Scorecard - a pilot study". *Journal of Business Ethics*, 55, pp. 173-186
- Von Bergen, C.W., Benco, D.C. (2004) "A Balanced Scorecard for Small Business". *Proceedings of the United States Association for Small Business and Entrepreneurship Conference*, pp. 1-21